

The
University
Of
Sheffield.

**Tropical legume trees and their
soil-mineral microbiome:
biogeochemistry and routes to
enhanced mineral access**

A thesis submitted
by **Dimitar Zdravkov Epihov**
in partial fulfilment of the requirements for the
Degree of Doctor of Philosophy in the
Department of Animal and Plant Sciences,
University of Sheffield

October 29th, 2018

© Copyright by Dimitar Zdravkov Epihov, 2018.

All rights reserved.

Acknowledgements

First and foremost, I would like to thank my academic supervising team including Professor David J. Beerling and Professor Jonathan R. Leake for their guidance and support throughout my PhD studies as well as the European Research Council (ERC) for funding my project.

Secondly, I would like to thank my girlfriend, Gabriela, my parents, Zdravko and Mariyana, and my grandmother Gina, for always believing in me. My biggest gratitude goes for my girlfriend for always putting up with working ridiculous hours and for helping me during field work even if it meant getting stuck in the Australian jungle at night and stumbling across a well-grown python.

I would also want to express my thanks to my first ever Biology teacher Mrs Moskova for inspiring and nurturing the interest that grew to be a life-lasting passion, curiosity and love towards all things living.

Lastly, I would like to thank Irene Johnson, our laboratory manager and senior technician for always been there for advice, help and general cheering up as well as all other great scientists and collaborators I have had the chance to talk to and work with during my PhD project.

I devote this work to a future with more green in it.

Dimitar

October, 2018

Introductory notes

The structure of this work is according to the the new alternative thesis format approved by the Faculty of Science at The University of Sheffield (more information can be found at <https://www.sheffield.ac.uk/rs/code/altformat>). The alternative format entails chapters be written as part of published or pre-published journal articles.

The Introduction chapter includes a paper titled “N₂-fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?” published in the peer-reviewed journal “Proceedings of the Royal Society: Biological Sciences” in July, 2017 (free-access to the article is available online at the Royal Society website: <http://rspb.royalsocietypublishing.org/content/284/1860/20170370>). Chapters 1, 2 and 3 as well as the Discussion & Conclusions chapter include pre-publication work formatted according to chosen target journals and their specific requirements.

All the laboratory experimental and computational work within this thesis was carried out by myself with the exception of fieldwork which included a certain amount of assistance from collaborators and my partner. The draft original version of each the five chapters were written by me with co-authors contributing to subsequent revisions during the editing process.

To summarise the contents of this thesis briefly: the introduction discusses the evolution of the first legume-rich tropical forests in the early Cenozoic (58-42 Mya) and the effect they had on biogeochemical cycles in general and silicate rock weathering and climate in particular. This work puts forward a hypothesis that early Cenozoic forests rich in N₂-fixing legumes caused enhanced weathering regimes globally, stimulating the drawdown of atmospheric CO₂.

Chapter 1 utilises a large-scale replicated weathering field study in Neotropical secondary forests in Panama rich in fixing legume trees demonstrating that N₂-fixers exerted 2-fold greater silicate rock weathering than non-fixers linked to significant differences in soil acidity and belowground microbial community structure and function.

Chapter 2 summarises our findings from another tropical system along a secondary forest chronosequence in the Australian Wet Tropics, showing that N₂-fixing and ectomycorrhizal monodominant *Acacia celsa* drives greater P and K-specific weathering dissolution from basalt silicate rocks throughout the chronosequence. Highly nodulated *Acacia* trees were also linked to greater total basalt and dunite weathering rates than non-fixers. Analyses from chapters 1 and 2 employ Next-Generation sequencing technology and omics-driven approach in demonstrating a consistent effect of N₂-fixing legume trees on their belowground microbiome in tropical forests that is that high inputs of fixed N are a strong determinant in functionally entraining the microbial community to high levels of mineral weathering. Consequently, this work also highlighted target candidate microbial genes and metabolic pathways linking those differences to enhanced weathering.

In order to test some of the highlighted candidate weathering genes *in vitro*, the next study, described in detail in Chapter 3 deployed the large transposon-mutant collection of the tropical soil β -proteobacterium *Burkholderia thailandensis* E264, selecting 11 mutant lines for further analysis. Data from replicated *in vitro* weathering experiments implicated several genes in the process of bacteria-mediated weathering. Further phylogenetic and taxonomic analyses of the soil pools of these genes indicated many elusive and uncultured bacterial lineages as carriers of those weathering genes thus highlighting their potential role in soil mineral weathering. This included the proposition and description of a new class within the Acidobacteria phylum, *Ca. Acidipotentia*, cl. nov.

The final Discussion and Conclusion chapter covers meta-analyses of data from my original datasets as well as already published literature to establish the role of N₂-fixers in ecosystem succession beyond their ecosystem N enrichment effects. The results from this chapter provide evidence for enhancements in acidification, soil lithotrophy, microbial respiration and anaerobic metabolism in soil beneath N₂-fixers that may all converge in a second previously unrecognised ecosystem service carried by N₂-fixers: that of enhanced weathering and subsequent increase in available nutrient stocks in such early successional systems.

Thesis Contributions Note

All the laboratory experimental and computational work within this thesis was carried out by myself with the exception of fieldwork which included a certain amount of assistance from collaborators and my partner. The draft original version of each the five chapters were written by me with co-authors contributing to subsequent revisions during the editing process. A chapter-specific breakdown:

Chapter 1 – D.Z.E. wrote the draft, carried out the analyses and conceived the hypothesis; all co-authors contributed to edits and revisions of the main manuscript.

Chapter 2 – D.Z.E. carried out the field and laboratory work and wrote the first draft; D.Z.E., D.J.B. and J.R.L. designed the experimental design; all co-authors contributed to edits and revisions of the main manuscript.

Chapter 3 – D.Z.E. and A.W.C. carried out the field work; D.Z.E. carried out the laboratory work and wrote the first draft; D.Z.E., D.J.B., A.W.C. and J.R.L. came up with the experimental design; all co-authors contributed to edits and revisions of the main manuscript.


Chapter 4 – D.Z.E. designed the experiment and carried out the experimental and statistical work related to it; D.Z.E wrote the first draft and D.J.B. and J.R.L have contributed to it with edits and comments.

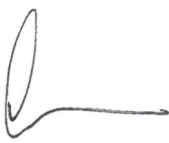
Chapter 5 – D.Z.E. wrote the first draft and carried out the meta-analyses found within it; all co-authors contributed with comments.

Signatures of co-authors:

Dimitar Z. Epihov 

David J. Beerling 

Jonathan R. Leake 

Alexander W. Cheesman 

Sarah A. Batterman 

Lars O. Hedin 

Kristin Saltonstall 

Jefferson S. Hall 

Lucas A. Cernusak 

Lisa M. Smith 

Michael I. Bird 

Susan G. W. Laurance 

Contents

1. Introduction: N ₂ -fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?	7
(Main Text)	8
(Figures and Figure Text)	30
(Supplementary Information and Figures).....	35
2. Chapter 1: Legume-microbial interactions unlock mineral nutrients during tropical forest succession.....	40
(Main Text).....	41
(Figures and Figure Text).....	51
(Supplementary Information and Figures).....	56
3. Chapter 2: Multiple symbiont recruitments and nitrogen feedbacks govern enhanced mineral access in monodominant <i>Acacia</i> of Australian tropical forests....	64
(Main Text).....	65
(Figures and Figure Text).....	82
(Supplementary Information and Figures).....	90
4. Chapter 3: Metabolic and genetic basis of bacteria-driven mineral weathering and its phylogenetic diversity at the soil-mineral interface.....	97
(Main Text).....	98
(Figures and Figure Text).....	118
5. Discussion and Conclusion: The biogeochemistry and belowground metagenomics of N ₂ -fixers: in search of fixer effects beyond nitrogen enrichment.....	123
(Main Text).....	124
(Figures and Figure Text).....	139
(End).....	147

Introduction

N₂-fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?

Dimitar Z. Epihov¹, Sarah A. Batterman², Lars O. Hedin³, Jonathan R. Leake¹, Lisa M. Smith¹ and David J. Beerling¹

¹Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK

²School of Geography and Priestley International Centre for Climate, University of Leeds, Leeds LS2 9JT, UK

³Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544, USA

Keywords: tropical forests, N₂-fixation, legume trees, rock weathering, CO₂ sequestration, Cenozoic

Fossil and phylogenetic evidence indicates legume-rich modern tropical forests replaced Late Cretaceous palm-dominated tropical forests across four continents during the early Cenozoic (58-42 Myr ago). Tropical legume trees can transform ecosystems via their ability to fix atmospheric N₂ and higher leaf N compared to non-legumes (35-65%) but it is unclear how their evolutionary rise contributed to silicate weathering, the long-term sink for atmospheric CO₂. Here we hypothesize that the increasing abundance of N₂-fixing legumes in tropical forests amplified silicate weathering rates by increased input of fixed N to terrestrial ecosystems via interrelated mechanisms including increasing microbial respiration and soil acidification, and stimulating forest net primary productivity. We suggest the high CO₂ early Cenozoic atmosphere further amplified legume weathering. Evolution of legumes with high weathering rates was likely driven by their high demand for phosphorus and micronutrients required for N₂-fixation and nodule formation.

1. Introduction

Biogeochemical weathering of silicate rocks (e.g., basalt, andesite, dunite) is a key process in the carbon cycle that acts as a long-term sink of atmospheric CO₂ [1]. Consumption of CO₂ by weathering is small (0.10-0.12 Gt C yr⁻¹) on an annual basis [2] compared to carbon transfers in photosynthesis or respiration. However, net CO₂ consumption by weathering is the dominant sink in the global carbon balance thus controlling atmospheric CO₂ and climate patterns at scales of millennia or longer [2].

Numerous field studies have shown that plants accelerate rock weathering through a suite of increasingly well understood processes [3] (Supplementary Figure 1). By increasing the soil pools of H⁺ ions, carbonic (H₂CO₃, from plant or soil respiration) and chelating organic (RCOO⁻) acids, plants and their symbiotic partners cause the weathering release of base cations (Supplementary Figure 1) that ultimately lead to the formation of marine carbonates on the seafloor [2]. The rise of the first forests during the Devonian period (416-359 Myr ago) [4] likely accelerated silicate weathering, contributing to the drawdown of atmospheric CO₂ and establishing the basic features of the modern land carbon cycle. Today, forests are thought to enhance rock weathering by a factor of 2-10 compared to unvegetated catchments [5].

During the Cenozoic (last 65 million years), the global biome transformation from palm-dominated late Cretaceous forests to the highly productive and carbon-rich tropical forests that exist today, discussed in more detail in the next section, included the rise of trees in the ecologically important legume family (Leguminosae, or “legumes”). Legumes dominate large areas of modern tropical forests in both total number of tree species and in abundance within local forests [6].

Four lines of evidence suggest that the evolution of the N₂-fixing rhizobial symbiosis (in which dinitrogen-fixing rhizobial bacteria are housed within specialized root nodules [7]) occurred as legumes radiated and spread in the early Cenozoic [8]. First, a whole-genome duplication event in the Papilionoideae clade, molecularly dated to 58 Myr ago, likely created the gene copies necessary for nodulation and N₂ fixation to evolve [9]. Second, many modern rainforest N₂-fixing legume trees are nodulated by β -rhizobia in the *Burkholderia* group [10]. Horizontal transfer of symbiotic *nod* genes between α -rhizobia and South American *Burkholderia* is dated to 60-50 Myr ago [11], indicating that compatible N₂-fixing host trees may have appeared at that time. Third, this molecular evidence is further strengthened by the presence of fossil legume genera recorded in early Cenozoic deposits with present-day relatives that are capable of symbiotic N₂ fixation (Supplementary Figure 2), with our synthesis indicating that the majority of fossil taxa identified at the genus level and recovered from the Palaeocene and Eocene belong to N₂-fixing genera (25 taxa) relative to non-fixing (16 taxa). Fourth, early Cenozoic sites in the Bighorn Basin show an increased proportion of legume fossil leaves recovered from 56 Myr old strata correlating with intensification of insect damage, a pattern consistent with the influx of fresh, fixed N into the ecosystem [12]. Fossil genera, the symbiotic status of their nearest living relatives (Supplementary Figure 2), evidence of increased insect damage in the fossil record in likely response to high foliar N and molecular clock dating therefore appear to indicate that N₂-fixation and diverse mycorrhizal symbioses have evolved in legumes by the early Cenozoic. Here, we review the rise of N₂-fixing legume-rich tropical forests early in the Cenozoic and propose a new testable hypothesis for how the evolution of this biome may have strengthened the long-term carbon cycle feedbacks that helped shape Earth's CO₂ and climate history in the Cenozoic and today.

2. Global rise of nitrogen-fixing legume-rich tropical forests

Late Cretaceous tropical floras were dominated by widely distributed palm communities from Africa to South America, a floristic region known as the Palmae Province [13],[14]. Communities in both the Paleo- and Neotropics contained abundant palms, including those similar to extant *Nypa* palms and suggestive of coastal intertidal habitats similar to mangrove forests, while other areas harboured palm-dominated dry forest communities. Unlike modern tropical forests, both of these communities were deprived of abundant dicot arboreal flora [13],[14]. In Africa, leaf fossil and pollen evidence indicate that the dominant palm lineages began to decline around the Cretaceous-Paleogene boundary [15] and completely disappeared in the fossil record after the late Miocene (23 Myr ago)[14]. Similarly, palm abundance in Neotropical areas decreased in the early Cenozoic, although palms remain an important element of these forests today [16]. The Palmae Province was replaced in Africa and assimilated in South America by the rise of modern tropical forests during the early Cenozoic. The earliest record of modern Neotropical forests – found in Colombia and dated to the late Palaeocene (58 Myr ago) – indicates that the flora resembled the current day composition of plant families with abundant fossilized dicot and palm leaves, including numerous legumes [17]. Pollen records from Africa similarly show the rise of modern families of dicot trees following the Palaeocene [14],[15].

Pollen and leaf macrofossils indicate that legume taxa have comprised a key component of tropical forests since the early Cenozoic (Figure 1a, Supplementary Figure 2). While it is difficult to translate a taxon's abundance in the fossil record to abundance in a forest, the persistent recovery of legume pollen, leaves, flowers, fruits and wood indicate that legume trees were present and widespread in the flora of the Americas and Africa. The following observations can be drawn from early Cenozoic records: (1) legume leaves made up 21-73% of all fossilized leaves in South and North American forest assemblages[18],[19]; (2) legumes

comprised 14-33% of all recorded taxa across tropical forests (Figure 1)[20],[21],[22], (3) single legume tree species represented up to 7% of all fossil leaves (>200 leaves) in species-diverse South American dry forests[23],[24]; (4) one legume tree species (the non-fixing *Cynometra*) formed a monodominant forest in Africa 46 Myr ago [21], with further monodominance indicated by the presence of Eocene fossils that belong to modern monodominant genera such as the non-fixing *Brachystegia* and *Julbernardia* (Eurasian deposits) and the non-fixing *Peltogyne* (South American formations) (Supplementary Figure 2); (5) rainforests with abundant presence of caesalpinoid and mimosoid (many modern representatives of which are N₂-fixing [25]) legumes were recorded in central Africa [14]; and, (6) tropical and temperate N₂-fixing legume trees may have coexisted during warm Eocene climates in higher latitude boreotropical forests (England, Hungary, North America)[26].

Fossil evidence, therefore, indicates that early Cenozoic tropical forests (wet, dry and boreotropical) had evolved abundant legumes across continents as well as latitudes (Figure 1a). The timing of the early Cenozoic assembly of legume-rich tropical forests (58 - 42 Myr ago) as documented by the fossil record is similar to the molecular clock dated diversification events in the legume clade (Figure 1b; for recent changes in legume taxonomy see [28]). Beneath these emerging tropical forests were substantial areas of unweathered rocks in tropical India [29], in South America, including the southeast part of the Amazon basin, and in the Amazon deltaic area [30] coinciding with peaks in terrestrial weathering (Figure 1b) as evident from the recovery of highly weathered palaeosols [31].

3. Mechanisms of N₂-fixing legume-driven enhanced weathering

Here, we propose that the rise of N₂-fixing legume trees enhanced weathering through a series of processes associated with three abilities especially well developed in this group of trees:

(1) to fix atmospheric N_2 , (2) to build disproportionately N-rich leaf tissue, and (3) to stimulate the primary production in ecosystems by redistributing fixed nitrogen to the soil and to neighboring trees.

First, nitrogen-fixing legumes have the ability to fix nitrogen at high rates in natural ecosystems [32]. Over time, fixers bring in substantial quantities of nitrogen and can provide the largest natural source of new nitrogen to ecosystems [33]. Soil nitrogen is high and nitrate and denitrification losses large (exceeding or rivaling many temperate forests exposed to nitrogen deposition) in tropical forests that harbor nitrogen fixers [32]. In a survey across 55 tropical forests, these systems naturally sustained loss rates of 4-6 kg N/ha nitrate, 6-10 kg N/ha of total dissolved nitrogen, and 4-5 kg N denitrified; when corrected for low levels of atmospheric nitrogen deposition, these rates could only be explained by fixation [34].

Second, N_2 -fixing legumes contain substantially higher leaf nitrogen than non-fixing tree species [35]. We performed a meta-analysis of 31 studies encompassing 561 tropical tree species ($n = 680$ measurements) to evaluate the nitrogen content of N_2 -fixing and non-fixing trees in natural forests and plantations across 22 different tropical regions (Figure 2a,b). Our analysis shows that, despite considerable variation across sites, N_2 -fixers exhibit higher mean leaf N content than non-fixers (by 35% in natural tropical forests and by 65% in tropical forestry plantations) and non-fixing legumes (by 21%). These findings are consistent with a study of leaf nitrogen across Amazonian tropical forests that also reported nitrogen-fixing legumes had higher leaf nitrogen content than both non-fixers as a whole and non-fixing legumes [36].

Third, this nitrogen-rich leaf tissue would cause increased input of N-rich compounds including proteins and amino acids to soils via litterfall. Such increased N input, in turn, would enrich soils in N and likely cause higher rates of productivity for non-fixing as well as N_2 -fixing trees. Evidence for such a major ecosystem impact comes from recent field studies:

N₂-fixing legumes provided ~50% of the nitrogen required for early growth of Panamanian secondary rainforests, supported rapid carbon accumulation in biomass of both fixers and non-fixers [37] and enhanced soil nitrogen [38] during periods of nitrogen limitation. Levels of N₂-fixation in early Cenozoic fixers are hard to establish empirically but indirect evidence of greater insect damage from fossil leaves together with greater palatability and protein content of N₂-fixing trees [12] support the assumption that ancient N₂-fixers were capable of generating high N foliage.

We suggest these three characteristics of N₂-fixing legumes likely entrain a suite of direct and indirect mechanisms that can enhance rates of rock weathering, as discussed below.

(a) N₂-fixing legume litter decomposition and microbial respiration

Litterfall and the decomposition of protein-enriched biomass would ultimately increase the flux of new fixed N into several linked soil processes (soil respiration, ammonification, nitrification) and pools (soil organic matter, dissolved organic nitrogen). The input of new nitrogen, in turn, would trigger several weathering-related mechanisms (Figure 2d).

First, the low C/N ratio of N₂-fixing legume litter implies fast decomposition, greater microbial respiration, and greater CO₂ production than non-legume litter [39],[40]. During decomposition, the majority of N-rich leaf tissue and its amino acids, amino sugars and other N-rich monomers will undergo ammonification and nitrification. Decomposition also generates organic acids and faster decomposition rates may facilitate passing the organic acid concentration threshold necessary to drive mineral weathering [41].

Second, N-rich organic matter can itself stimulate soil microbial activity and respiration. Although C inputs would have similar effects regardless of whether derived from decomposition of leguminous N-rich or non-leguminous N-poor litter, the lack of sufficient N can ultimately downregulate microbial respiration specifically under high CO₂ regimes

[42], such as those seen during the early Cenozoic (Figure 1b). Addition of N_2 -fixing legume-derived N-rich litter may therefore have a dual function. First, it will fuel microbial respiration with the energy stored in the C-H and C-C bonds of its carbohydrate component. Second, because of its abundance in N and protein, it will promote microbial respiration by alleviating any existing N limitation over microbial metabolism. *In situ* studies in tropical soils confirm augmented rates of microbial respiration in the combined glucose and N treatment compared to the glucose treatment alone [43].

Third, the dissolved CO_2 generated by microbial respiration forms carbonic acid (H_2CO_3) which, in turn, acts as a major weathering agent [44](Supplementary Figure 1). Increased microbial respiration also positively correlates with the production of chelating organic acids, e.g., gluconic acid, a secreted by-product of microbial catabolism [45].

(b) N_2 -fixing legume-driven soil acidification

Ammonia generated by ammonification during litter decomposition can undergo nitrification. In the process, each molecule of ammonia converted to nitrate generates three by-product H^+ ions. Although these H^+ ions are typically counterbalanced by plant secretion of anions (bicarbonate or organic acids) for each acquired NO_3^- , nitrate leaching can uncouple this relationship and promote the buildup of H^+ in the soil. High levels of N_2 -fixation can exceed the rates at which N is immobilized within the system, resulting in enhanced NO_3^- leaching (as discussed above) and enhanced transport of H^+ to deeper soil horizons (where contents of unweathered minerals may be high). Tree ring data from tropical fossil woods indicate that climate seasonality was largely similar between early Cenozoic and modern tropical forests [46] supporting the view that nitrification patterns as affected by soil moisture/dryness [47] likely were comparable.

During the leaching of NO_3^- large amounts of counterbalancing cations (Ca^{2+} , Mg^{2+} , K^+) released by cation exchange reactions with nitrification-generated H^+ are leached too,

resulting in the decline of soil cation exchange capacity and soil pH buffering capacities. This phenomenon has been recorded for N₂-fixing forests of *Alnus rubra* in which large inputs of fixed N caused leaching, decreased cation concentration and increased soil acidification [48]. Despite the tight N budget of most tropical forest systems, substantial levels of nitrate leaching still occurs [32] suggesting that similar mechanisms likely operate in tropical forests rich in N₂-fixing legumes. In addition, because of their N₂-fixation, fixers tend to acquire lower relative amounts of negatively charged ions such as NO₃⁻ and subsequently extrude greater amounts of H⁺ to balance increased internal relative concentration of positively charged ions [49].

Consequently, pronounced soil acidification has been recorded in various N₂-fixing species from herbs [49],[50] to trees and shrubs of temperate forest [51],[52] and tropical rainforest [38] areas. Recent analysis of tropical rainforests at four Neotropical locations revealed that forests rich in N₂-fixers exhibited increased soil acidity (pH 4.1) and lower Ca²⁺ and Mg²⁺ concentrations than forests poor in N₂-fixing legumes (pH 5.2) [53]. N₂-fixing legume-driven acidification can promote weathering by acid attack (acidolysis) of the mineral lattice (Supplementary Figure 1) but also by depleting soil cations through cation exchange, thus shifting the equilibrium towards further mineral dissolution.

(c) N₂-fixing legume-driven stimulation of net primary productivity (NPP)

Ultimately, inorganic forms of fixed N are acquired from the soil solution by roots stimulating the N input into biomass, including that of neighboring non-fixing trees. For instance, the non-fixing tropical trees *Peschiera*, *Psidium* [54], *Eucalyptus* [55] and *Terminalia* [56] all exhibited increased foliar N levels in N₂-fixing legume-rich neighbourhoods compared to legume-poor settings. As foliar N correlates with increased levels of crude leaf protein, including the photosynthetic enzyme RUBISCO [57], the photosynthetic rates of individual

trees and the NPP of such mixed fixer/non-fixer forests may be up-regulated. Indeed, N₂-fixing legumes exhibit up to 2-fold greater photosynthetic rates than the less N-rich leaves of non-fixing trees in Zimbabwe [58]. Similarly, non-fertilized mixed non-fixer/N₂-fixer forestry plantations reveal augmented NPP rates compared to non-fixing forests in Brazil and Puerto Rico [59],[60].

Fossil evidence supports N₂-fixing legume-driven N-fertilization on productivity of tropical ecosystems. Presumed nitrogen-fixing legume-dominated assemblages exhibited insect damage (linked to higher leaf nitrogen content) spread across fossil taxa relative to systems with fewer legumes in which foliar damage was more concentrated on legume leaves [12]. This observation indicates that as legume domination was established, N redistribution triggered by the input of N-rich litter increased N levels of neighbouring non-legumes (as observed in modern systems). The source of this N buffering effect is better explained by legumes capable of N₂-fixation than non-fixing legumes because the patterns are consistent with the influx of new fixed N to the system.

Given that some canopy photosynthate from highly productive N₂-fixing legume-rich forests will be allocated to symbiotic mycorrhizal fungi, the mycelial networks of these fungi that grow in intimate contact with mineral grains, may drive enhanced rock weathering and inorganic nutrient release via chelation, carbonation and acidolysis (Supplementary Figure 1)[63]. Greater gross primary production (GPP) and its related NPP rates also correlate with greater root respiration (with associated production of carbonic acid) and organic acid leaching, which promotes further weathering [63] (Supplementary Figure 1). N₂-fixing legume-enhanced forest NPP can also increase the demand for nutrients and thus further necessitate more extensive soil exploration via roots and mycorrhizal fungi and eventually enhanced rock weathering. Therefore, increased N inputs could indirectly increase rock

weathering via stimulation of rainforest NPP in legume-rich communities compared to *Nypa* and other late Cretaceous palm forests as well as to legume-poor early Cenozoic analogues.

(d) Accessory mechanisms of N₂-fixing legume-driven weathering

The unique ability of legumes (including many rainforest N₂-fixing legume trees [64],[65]) to synthesize and exude isoflavonoids [66] may also have some impact on weathering rates. Isoflavonoids enhance P and Fe solubilization from the mineral vivianite by acting as soil chelators (Supplementary Figure 1) as well as by decreasing organic acid decomposition [67]. Comparison between the estimated low-molecular organic acid exudation by lowland tropical rainforest trees (~25 µg C g⁻¹ DW root h⁻¹) [68] and isoflavonoid exudation of the N₂-fixer *Lupinus albus* (~31 µg C g⁻¹ DW root h⁻¹) [69] (see Supplementary Information for detailed calculations) suggests that isoflavonoids could contribute to the pool of plant-derived chelating agents in legume-rich forest soils.

Isoflavonoids are crucial in establishing the N₂-fixing legume-rhizobial symbiosis by enabling both attraction and priming of rhizobial partners [70]. They attract larger soil rhizobial populations [71] of nodulation-competent strains of *Burkholderia*, *Rhizobium*, and *Mesorhizobium* – members of all of these genera have been shown to exert strong chelating activities [72]. Soil pH, C, N and C/N ratio are also important determinants of microbial community structure [73]. Finally, legume-mediated changes in soil chemistry may change microbial community of the mineralosphere selecting for nitrophilic and acidophilic bacterial taxa.

4. N₂-fixing legume-rich forest responses to a CO₂-rich early Cenozoic atmosphere

The rise of N₂-fixing legume-rich tropical forests during the early Cenozoic coincides with elevated atmospheric CO₂ concentrations, with potential feedbacks on primary production and weathering (Figure 1b, Figure 3). Evidence for the mechanisms that may govern this potential feedback comes from Free Air CO₂ Enrichment (FACE) experiments. In the Oak Ridge, Tennessee, FACE experiment the non-fixing, AM *Liquidambar styraciflua* trees showed a 24% increase in NPP during the first 6 years of exposure to elevated CO₂ [74]. However, over the next 5 years the positive CO₂ enrichment effect decreased to +9% in 11-year old stands as ecosystem N stocks declined [74], suggesting progressive soil N-limitation on tree NPP in the long-term under high CO₂ [74],[75]. N₂-fixing legumes may mitigate this N-limitation mechanism under a high CO₂ atmosphere because N-limitation would favour recruitment of N₂-fixing legumes and/or upregulate their fixation rates [76],[77]. Fossil evidence suggests that N₂-fixing legumes may increase in abundance under such conditions. During the transient climate warming event across the Palaeocene-Eocene Thermal Maximum (PETM; 55.8 Myr ago) that is linked to a rise in atmospheric CO₂ and continental weathering regimes [78], the abundance of fossilized leguminous leaf specimens increased to 73% and then declined to 21% post-PETM in the Bighorn Basin, US [20]. Further evidence from PETM sites dominated by legumes corroborates extensive N₂-fixation capacity increasing nitrogen availability to the system (as discussed above) [12].

Physiologically, elevated CO₂ can promote nodulation and N₂-fixation [79],[80],[81], mycorrhization [82] and photosynthetic rates, and therefore may allow N₂-fixing legume productivity to increase proportionally more in response to CO₂ than non-legumes [79],[81]. Furthermore, nodules represent additional sinks exchanging the increased flux of assimilates for fixed N thus curtailing the photosynthetic acclimation to elevated CO₂ when unconstrained by other factors [83], allowing higher photosynthetic rates to persist. Those effects could promote N₂-fixer recruitment, up-regulated N₂-fixation rates and greater dominance at high

CO₂ concentrations [84]. A FACE experiment at Oak Ridge analyzed the CO₂ response of over 2000 seedlings from 14 different temperate tree species. After 5 years, the N₂-fixing legume *Robinia pseudoacacia* exhibited an order of magnitude higher biomass response than all of the non-fixing angiosperm trees [85]. Controlled environment pot-based CO₂-enrichment experiments indicate that the photosynthesis and growth responses of nodulated N₂-fixing Leguminosae rainforest trees were significantly greater than that of non-leguminous species investigated [86]. Although there are clear limitations in extrapolating from these studies to legumes of early Cenozoic tropical forests, the mechanistic basis of the CO₂ response – linked to alleviation of N-limitation – would still hold.

Based on these findings, we conceptualize that different feedback loops operated between non-legume and N₂-fixing legume forests, atmospheric CO₂ and climate in the Cenozoic (Figure 3). In *non-fixing forests* like those that existed prior to legume evolution or in legume-poor tropical forests of the early Cenozoic, increased atmospheric CO₂ would stimulate NPP until available soil resources – likely N and P in many locations – are exhausted (Figure 3a). Progressive N-limitation could therefore uncouple the ‘standard’ relationship between NPP, CO₂ and weathering [87] in non-legume forests. In contrast, however, in *legume-rich forests*, progressive N limitation would likely further promote recruitment of N₂ fixers and the up-regulation of N₂ fixation rates (Figure 3a), as observed in modern N-limited rainforests [37]. This could allow NPP to respond to increasing CO₂ and help promote continued weathering (Figure 3b). Additionally, biological weathering processes are strengthened by inputs of N-rich legume litter, and associated downstream processes. Combined this evidence indicate that in CO₂-rich conditions, the significant role of legumes in maintaining enhanced weathering regimes in early tropical forests may be amplified.

5. Evolutionary drivers of enhanced weathering by N₂-fixing legumes

Central to our feedback analyses (Figure 3) is the idea that N₂-fixing legumes are associated with higher weathering rates than non-legume trees. This effect, in turn, may have evolved in response to a disproportionately high demand for phosphorus (P), molybdenum (Mo) and iron (Fe) across legume taxa. P and Mo have been identified as potentially limiting factors of N₂ fixation within tropical forests [88],[89]. These limitations may occur because the most common type of nitrogenases involved in symbiotic N₂-fixation requires a Fe/Mo complex acting as a cofactor [90] while high P intake accommodates for enhanced production of energy-rich metabolites (*e.g.*, ATP) and membranes during nodule organogenesis [91]. Linked to the likely greater P demand driven by higher rates of growth, some but not all N₂-fixing legumes may have higher foliar P levels than non-fixing trees (Supplementary Table 1). Fe is also required for production of leghaemoglobin in nodules for oxygen binding [92]. Fe is very abundant in tropical soils but it is highly insoluble. Most P in soils is also insoluble in complexes with Al- and Fe-bearing secondary minerals, and fresh Mo and P inputs originate from weathering of otherwise plant-unavailable mineral sources. Both the dissolution of insoluble P and Fe and the release of mineral-bound Mo rely upon the same weathering mechanisms that include chelation and acidolysis [93] (Supplementary Figure 1). Aluminium and iron phosphate minerals such as variscite and vivianite, respectively, dissolve faster at pH<6, a process exacerbated by organic acids[93],[94].

Overall, the processes of N₂ fixation and nodule formation require an array of sparingly soluble (P, Fe) or scarce soil minerals (Mo). This observation suggests that the mechanisms of enhanced weathering overlap with those driving acquisition of elements essential for N₂-fixing legumes. It provides a mechanism that would promote the evolution of adaptive strategies in tropical legumes leading to enhanced weathering and thereby unlocking sparingly soluble limiting nutrients. Our hypothesized mechanisms that relate N₂-fixing legume functioning to weathering rates are suitable for direct investigation in the field and lab and

future studies will hopefully further elucidate the relative importance of each of the mechanisms of the hereby proposed hypothesis.

6. Conclusions

Fossils and molecular dating suggest that a worldwide shift from palm-dominated communities to ‘modern’ tropical forests occurred early in the Cenozoic and involved the development of N₂-fixing legume-rich and symbiotically diverse communities. Based on our analyses of potential effects on forest ecosystem biogeochemical C and N cycling, we propose that the increasing abundance of N₂-fixing legumes in tropical forests amplified weathering rates through several interconnected pathways. Firstly, N₂-fixing legumes increased soil inputs of N-rich organic matter (by an estimated 35-65% based on modern analogues) which can promote microbial respiration and carbonation as well as progressive soil acidification resulting from leaching and compensatory H⁺ extrusion. Subsequently, increased N inputs may have fuelled greater N availability stimulating forest NPP thus driving further carbonation, organic acid chelation and rhizospheric weathering activities. Lastly, exudation of isoflavonoids unique to legumes could have provided an additional source of chelating activities that cause rock weathering. Together with soil acidification and decreasing C/N ratios these effects could have indirectly driven shifts in the weathering-potential of the soil microbial community.

We suggest the global evolution of tropical forests rich in N₂-fixing legumes in the early Cenozoic in concert with abiotic drivers, including reduced subduction of oceanic crust and the rise of the Himalayas/Tibetan plateau [29],[95], could have driven enhanced weathering regimes over large pantropical areas with consequent feedbacks on global climate stabilization. Furthermore, N₂-fixing legumes help maintain the NPP response to atmospheric CO₂ concentration. In an evolutionary context, tropical N₂-fixing legumes appear to enhance

CHANGED!

rock weathering as a possible adaptation to unlock previously unavailable P, Mo and Fe mineral sources, thus alleviating limitations on N₂ fixation processes.

Authors' contributions. D.Z.E., S.B., L.H. and D.J.B conceived the review; D.E. compiled the first draft; all authors contributed to revisions of the paper.

Funding. D.Z.E. is supported by an ERC Advanced Grant awarded to D.J.B. (CDREG, 322998). S.B. was supported by a U.K. Natural Environment Research Council (grant NE/M019497/1) and a Climate Mitigation Initiative Young Investigator Fellowship (with funding from BP).

References

1. Pagani, M., Caldeira, K., Berner, R. and Beerling, D.J., 2009. The role of terrestrial plants in limiting atmospheric CO₂ decline over the past 24 million years. *Nature*, 460(7251), p.85.
2. Goudie, A.S. and Viles, H.A., 2012. Weathering and the global carbon cycle: Geomorphological perspectives. *Earth-Science Reviews*, 113(1-2), pp.59-71.
3. Taylor, L.L., Leake, J.R., Quirk, J., Hardy, K., Banwart, S.A. and Beerling, D.J., 2009. Biological weathering and the long-term carbon cycle: integrating mycorrhizal evolution and function into the current paradigm. *Geobiology*, 7(2), pp.171-191.
4. Algeo, T.J. and Scheckler, S.E., 1998. Terrestrial-marine teleconnections in the Devonian: links between the evolution of land plants, weathering processes, and marine anoxic events. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 353(1365), pp.113-130.
5. Moulton, K.L., West, J. and Berner, R.A., 2000. Solute flux and mineral mass balance approaches to the quantification of plant effects on silicate weathering. *American Journal of Science*, 300(7), pp.539-570.
6. Ter Steege, H., Pitman, N.C., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.F., Prévost, M.F., Spichiger, R., Castellanos, H. and Von Hildebrand, P., 2006. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, 443(7110), p.444.
7. Dilworth, M.J., James, E.K., Sprent, J.I. and Newton, W.E. eds., 2008. *Nitrogen-fixing leguminous symbioses* (Vol. 7). Springer Science & Business Media.
8. Lavin, M., Herendeen, P.S. and Wojciechowski, M.F., 2005. Evolutionary rates analysis of Leguminosae implicates a rapid diversification of lineages during the tertiary. *Systematic biology*, 54(4), pp.575-594.
9. Werner, G.D., Cornwell, W.K., Sprent, J.I., Kattge, J. and Kiers, E.T., 2014. A single evolutionary innovation drives the deep evolution of symbiotic N₂-fixation in angiosperms. *Nature Communications*, 5, p.4087.
10. Barrett, C.F. and Parker, M.A., 2005. Prevalence of Burkholderia sp. nodule symbionts on four mimosoid legumes from Barro Colorado Island, Panama. *Systematic and Applied Microbiology*, 28(1), pp.57-65.
11. Walker, R., Agapakis, C.M., Watkin, E. and Hirsch, A.M., 2015. Symbiotic nitrogen fixation in

legumes: perspectives on the diversity and evolution of nodulation by *Rhizobium* and *Burkholderia* species. *Biological nitrogen fixation*, 2, pp.913-923.

12. Currano, E.D., Laker, R., Flynn, A.G., Fogt, K.K., Stradtman, H. and Wing, S.L., 2016. Consequences of elevated temperature and pCO₂ on insect folivory at the ecosystem level: perspectives from the fossil record. *Ecology and evolution*, 6(13), pp.4318-4331.
13. Vajda, V. and Bercovici, A., 2012. Palynostratigraphy of the Cretaceous-Paleogene mass-extinction interval of the Southern Hemisphere. *Dicéngxué zázhi*, 36(2), pp.153-164.
14. Maley, J., 1996. The African rain forest—main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, 104, pp.31-73.
15. Jacobs, B.F., 2004. Palaeobotanical studies from tropical Africa: relevance to the evolution of forest, woodland and savannah biomes. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1450), pp.1573-1583.
16. Burnham, R.J. and Johnson, K.R., 2004. South American palaeobotany and the origins of neotropical rainforests. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, 359(1450), pp.1595-1610.
17. Wing, S.L., Herrera, F., Jaramillo, C.A., Gómez-Navarro, C., Wilf, P. and Labandeira, C.C., 2009. Late Paleocene fossils from the Cerrejón Formation, Colombia, are the earliest record of Neotropical rainforest. *Proceedings of the National Academy of Sciences*, 106(44), pp.18627-18632.
18. Jaramillo, C., Ochoa, D., Contreras, L., Pagani, M., Carvajal-Ortiz, H., Pratt, L.M., Krishnan, S., Cardona, A., Romero, M., Quiroz, L. and Rodriguez, G., 2010. Effects of rapid global warming at the Paleocene-Eocene boundary on neotropical vegetation. *Science*, 330(6006), pp.957-961.
19. Wing, S.L., Herrera, F., Jaramillo, C.A., Gómez-Navarro, C., Wilf, P. and Labandeira, C.C., 2009. Late Paleocene fossils from the Cerrejón Formation, Colombia, are the earliest record of Neotropical rainforest. *Proceedings of the National Academy of Sciences*, 106(44), pp.18627-18632.
20. Currano, E.D., Labandeira, C.C. and Wilf, P., 2010. Fossil insect folivory tracks paleotemperature for six million years. *Ecological Monographs*, 80(4), pp.547-567.
21. Cantrill, D.J., Bamford, M.K., Wagstaff, B.E. and Sauquet, H., 2013. Early Eocene fossil plants from the Mwadui kimberlite pipe, Tanzania. *Review of palaeobotany and palynology*, 196, pp.19-35.
22. Ettingshausen, C.B., 1879. III. Report on Phyto-Palaeontological investigations of the fossil Flora of Sheppey. *Proceedings of the Royal Society of London*, 29(196-199), pp.388-396.
23. Wilf, P., Johnson, K.R., Cuneo, N.R., Smith, M.E., Singer, B.S. and Gandolfo, M.A., 2005. Eocene plant diversity at Laguna del Hunco and Río Pichileufú, Patagonia, Argentina. *The American Naturalist*, 165(6), pp.634-650.
24. Barreda, V. and Palazzesi, L., 2010. Vegetation during the Eocene–Miocene interval in central Patagonia: a context of mammal evolution. *The paleontology of Gran Barranca: evolution and environmental change through the Middle Cenozoic of Patagonia*.
25. Sprent, J.I., 2009. *Legume nodulation: a global perspective*. John Wiley & Sons.
26. Doyle, J.J. and Luckow, M.A., 2003. The rest of the iceberg. Legume diversity and evolution in a phylogenetic context. *Plant physiology*, 131(3), pp.900-910.
27. Herendeen, P.S., 1992. The fossil history of the Leguminosae: phylogenetic and biogeographic implications. In *Advances in Legume Systematics, Part 4. The Fossil Record* (pp. 303-316). Royal Botanic Gardens.
28. Azani, N., Babineau, M., Bailey, C.D., Banks, H., Barbosa, A.R., Pinto, R.B., Boatwright, J.S.,

- Borges, L.M., Brown, G.K., Bruneau, A. and Candido, E., 2017. A new subfamily classification of the Leguminosae based on a taxonomically comprehensive phylogeny The Legume Phylogeny Working Group (LPWG). *Taxon*, 66(1), pp.44-77.
29. Jagoutz, O., Macdonald, F.A. and Royden, L., 2016. Low-latitude arc–continent collision as a driver for global cooling. *Proceedings of the National Academy of Sciences*, 113(18), pp.4935-4940.
30. Putzer, H., 1984. The geological evolution of the Amazon basin and its mineral resources. In *The Amazon* (pp. 15-46). Springer, Dordrecht.
31. Retallack, G.J., 2010. Lateritization and bauxitization events. *Economic Geology*, 105(3), pp.655-667.
32. Hedin, L.O., Brookshire, E.J., Menge, D.N. and Barron, A.R., 2009. The nitrogen paradox in tropical forest ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, 40, pp.613-635.
33. Vitousek, P.M., Menge, D.N., Reed, S.C. and Cleveland, C.C., 2013. Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1621), p.20130119.
34. Brookshire, E.J., Hedin, L.O., Newbold, J.D., Sigman, D.M. and Jackson, J.K., 2012. Sustained losses of bioavailable nitrogen from montane tropical forests. *Nature Geoscience*, 5(2), p.123.
35. Rascher, K.G., Hellmann, C., Máguas, C. and Werner, C., 2012. Community scale ¹⁵N isoscapes: tracing the spatial impact of an exotic N₂-fixing invader. *Ecology Letters*, 15(5), pp.484-491.
36. Fyllas, N.M., Patino, S., Baker, T.R., Bielefeld Nardoto, G., Martinelli, L.A., Quesada, C.A., Paiva, R., Schwarz, M., Horna, V., Mercado, L.M. and Santos, A., 2009. Basin-wide variations in foliar properties of Amazonian forest: phylogeny, soils and climate. *Biogeosciences*, 6, pp.2677-2708.
37. Batterman, S.A., Hedin, L.O., Van Breugel, M., Ransijn, J., Craven, D.J. and Hall, J.S., 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature*, 502(7470), p.224.
38. Shebitz, D.J. and Eaton, W., 2013. Forest structure, nutrients, and *Pentaclethra macroloba* growth after deforestation of Costa Rican lowland forests. *ISRN Ecology*, 2013.
39. Milcu, A., Partsch, S., Scherber, C., Weisser, W.W. and Scheu, S., 2008. Earthworms and legumes control litter decomposition in a plant diversity gradient. *Ecology*, 89(7), pp.1872-1882.
40. Schwendener, C.M., Lehmann, J., Rondon, M., Wandelli, E. and Fernandes, E., 2007. Soil mineral N dynamics beneath mixtures of leaves from legume and fruit trees in Central Amazonian multi-strata agroforests. *Acta Amazonica*, 37(3), pp.313-320.
41. Jones, D.L., Dennis, P.G., Owen, A.G. and Van Hees, P.A.W., 2003. Organic acid behavior in soils—misconceptions and knowledge gaps. *Plant and soil*, 248(1-2), pp.31-41.
42. Hu, S., Chapin III, F.S., Firestone, M.K., Field, C.B. and Chiariello, N.R., 2001. Nitrogen limitation of microbial decomposition in a grassland under elevated CO₂. *Nature*, 409(6817), p.188.
43. Ilstedt, U. and Singh, S., 2005. Nitrogen and phosphorus limitations of microbial respiration in a tropical phosphorus-fixing Acrisol (Ultisol) compared with organic compost. *Soil Biology and Biochemistry*, 37(7), pp.1407-1410.
44. Taylor, L.L., Banwart, S.A., Valdes, P.J., Leake, J.R. and Beerling, D.J., 2012. Evaluating the effects of terrestrial ecosystems, climate and carbon dioxide on weathering over geological time: a global-scale process-based approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1588), pp.565-582.
45. Velizarov, S. and Beschkov, V., 1998. Biotransformation of glucose to free gluconic acid by *Gluconobacter oxydans*: substrate and product inhibition situations. *Process Biochemistry*, 33(5), pp.527-534.
46. Wheeler, E.A. and Baas, P., 1991. A survey of the fossil record for Dicotyledonous wood and its

significance for evolutionary and ecological wood anatomy. *IAWA Journal*, 12(3), pp.275-318.

47. Birch, H.F., 1960. Nitrification in soils after different periods of dryness. *Plant and Soil*, 12(1), pp.81-96.
48. Van Miegroet, H. and Cole, D.W., 1984. The Impact of Nitrification on Soil Acidification and Cation Leaching in a Red Alder Ecosystem 1. *Journal of Environmental Quality*, 13(4), pp.586-590.
49. Raven, J.A., Franco, A.A., de JESUS, E.L. and Jacob-Neto, J., 1990. H⁺ extrusion and organic-acid synthesis in N₂-fixing symbioses involving vascular plants. *New Phytologist*, 114(3), pp.369-389.
50. Bolan, N.S., Hedley, M.J. and White, R.E., 1991. Processes of soil acidification during nitrogen cycling with emphasis on legume based pastures. *Plant and soil*, 134(1), pp.53-63.
51. Leary, J.K., Hue, N.V., Singleton, P.W. and Borthakur, D., 2006. The major features of an infestation by the invasive weed legume gorse (*Ulex europaeus*) on volcanic soils in Hawaii. *Biology and Fertility of Soils*, 42(3), pp.215-223.
52. Homann, P.S., Van Miegroet, H., Cole, D.W. and Wolfe, G.V., 1992. Cation distribution, cycling, and removal from mineral soil in Douglas-fir and red alder forests. *Biogeochemistry*, 16(2), pp.121-150.
53. Powers, J.S., Treseder, K.K. and Lerdau, M.T., 2005. Fine roots, arbuscular mycorrhizal hyphae and soil nutrients in four neotropical rain forests: patterns across large geographic distances. *New Phytologist*, 165(3), pp.913-921.
54. Laclau, J.P., Bouillet, J.P., Gonçalves, J.D.M., Silva, E.V., Jourdan, C., Cunha, M.C.S., Moreira, M.R., Saint-André, L., Maquère, V., Nouvellon, Y. and Ranger, J., 2008. Mixed-species plantations of *Acacia mangium* and *Eucalyptus grandis* in Brazil: 1. Growth dynamics and aboveground net primary production. *Forest Ecology and Management*, 255(12), pp.3905-3917.
55. Bini, D., Figueiredo, A.F., Silva, M.C.P.D., Vasconcellos, R.L.D.F. and Cardoso, E.J.B.N., 2013. Microbial biomass and activity in litter during the initial development of pure and mixed plantations of *Eucalyptus grandis* and *Acacia mangium*. *Revista Brasileira de Ciência do Solo*, 37(1), pp.76-85.
56. Nichols, J.D. and Carpenter, F.L., 2006. Interplanting *Inga edulis* yields nitrogen benefits to *Terminalia amazonia*. *Forest Ecology and Management*, 233(2-3), pp.344-351.
57. Evans, J.R., 1989. Photosynthesis and nitrogen relationships in leaves of C₃ plants. *Oecologia*, 78(1), pp.9-19.
58. Tuohy, J.M., Prior, J.A. and Stewart, G.R., 1991. Photosynthesis in relation to leaf nitrogen and phosphorus content in Zimbabwean trees. *Oecologia*, 88(3), pp.378-382.
59. Parrotta, J.A., 1999. Productivity, nutrient cycling, and succession in single-and mixed-species plantations of *Casuarina equisetifolia*, *Eucalyptus robusta*, and *Leucaena leucocephala* in Puerto Rico. *Forest Ecology and Management*, 124(1), pp.45-77.
60. Santos, F.M., de Carvalho Balieiro, F., dos Santos Ataíde, D.H., Diniz, A.R. and Chaer, G.M., 2016. Dynamics of aboveground biomass accumulation in monospecific and mixed-species plantations of *Eucalyptus* and *Acacia* on a Brazilian sandy soil. *Forest Ecology and Management*, 363, pp.86-97.
61. McKey, Doyle., 1994. Legumes and nitrogen: the evolutionary ecology of a nitrogen-demanding lifestyle. *Advances in legume systematics*, 5, pp.211-228.
62. Michalet, S., Rohr, J., Warshan, D., Bardon, C., Roggy, J.C., Domenach, A.M., Czarnes, S., Pommier, T., Combourieu, B., Guillaumaud, N. and Bellvert, F., 2013. Phytochemical analysis of mature tree root exudates in situ and their role in shaping soil microbial communities in relation to tree N-acquisition strategy. *Plant physiology and biochemistry*, 72, pp.169-177.
63. Taylor, L.L., Banwart, S.A., Valdes, P.J., Leake, J.R. and Beerling, D.J., 2012. Evaluating the effects of terrestrial ecosystems, climate and carbon dioxide on weathering over geological time: a global-scale

process-based approach. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 367(1588), pp.565-582.

64. Kraft, C., Jenett-Siems, K., Köhler, I., Siems, K., Abbiw, D., Bienzle, U. and Eicha, E., 2002. Andriol A and B, Two Unique 6-Hydroxymethylpterocarpenes from *Andira inermis*. *Zeitschrift für Naturforschung C*, 57(9-10), pp.785-790.

65. Leuner, O., Havlik, J., Hummelova, J., Prokudina, E., Novy, P. and Kokoska, L., 2013. Distribution of isoflavones and coumestrol in neglected tropical and subtropical legumes. *Journal of the Science of Food and Agriculture*, 93(3), pp.575-579.

66. Yu, O., Jung, W., Shi, J., Croes, R.A., Fader, G.M., McGonigle, B. and Odell, J.T., 2000. Production of the isoflavones genistein and daidzein in non-legume dicot and monocot tissues. *Plant physiology*, 124(2), pp.781-794.

67. Tomasi, N., Weisskopf, L., Renella, G., Landi, L., Pinton, R., Varanini, Z., Nannipieri, P., Torrent, J., Martinoia, E. and Cesco, S., 2008. Flavonoids of white lupin roots participate in phosphorus mobilization from soil. *Soil Biology and Biochemistry*, 40(7), pp.1971-1974.

68. Aoki, M., Fujii, K. and Kitayama, K., 2012. Environmental control of root exudation of low-molecular weight organic acids in tropical rainforests. *Ecosystems*, 15(7), pp.1194-1203.

69. Weisskopf, L., Tomasi, N., Santelia, D., Martinoia, E., Langlade, N.B., Tabacchi, R. and Abou-Mansour, E., 2006. Isoflavonoid exudation from white lupin roots is influenced by phosphate supply, root type and cluster-root stage. *New Phytologist*, 171(3), pp.657-668.

70. Oldroyd, G.E., 2013. Speak, friend, and enter: signalling systems that promote beneficial symbiotic associations in plants. *Nature Reviews Microbiology*, 11(4), p.252.

71. Odee, D.W., Sutherland, J.M., Kimiti, J.M. and Sprent, J.I., 1995. Natural rhizobial populations and nodulation status of woody legumes growing in diverse Kenyan conditions. *Plant and Soil*, 173(2), pp.211-224.

72. Uroz, S., Kelly, L.C., Turpault, M.P., Lepleux, C. and Frey-Klett, P., 2015. The mineralosphere concept: mineralogical control of the distribution and function of mineral-associated bacterial communities. *Trends in microbiology*, 23(12), pp.751-762.

73. Lauber, C.L., Hamady, M., Knight, R. and Fierer, N., 2009. Pyrosequencing-based assessment of soil pH as a predictor of soil bacterial community structure at the continental scale. *Appl. Environ. Microbiol.*, 75(15), pp.5111-5120.

74. Norby, R.J., Warren, J.M., Iversen, C.M., Medlyn, B.E. and McMurtrie, R.E., 2010. CO₂ enhancement of forest productivity constrained by limited nitrogen availability. *Proceedings of the National Academy of Sciences*, 107(45), pp.19368-19373.

75. LeBauer, D.S. and Treseder, K.K., 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, 89(2), pp.371-379.

76. Vitousek, P.M., Cassman, K.E.N., Cleveland, C., Crews, T., Field, C.B., Grimm, N.B., Howarth, R.W., Marino, R., Martinelli, L., Rastetter, E.B. and Sprent, J.I., 2002. Towards an ecological understanding of biological nitrogen fixation. In *The nitrogen cycle at regional to global scales* (pp. 1-45). Springer, Dordrecht.

77. Van Der Heijden, M.G., De Bruin, S., Luckerhoff, L., Van Logtestijn, R.S. and Schlaeppi, K., 2016. A widespread plant-fungal-bacterial symbiosis promotes plant biodiversity, plant nutrition and seedling recruitment. *The ISME journal*, 10(2), p.389.

78. Penman, D.E., 2016. Silicate weathering and North Atlantic silica burial during the Paleocene-Eocene thermal maximum. *Geology*, 44(9), pp.731-734.

79. Polley, H.W., Johnson, H.B. and Mayeux, H.S., 1997. Leaf physiology, production, water use, and nitrogen dynamics of the grassland invader *Acacia smallii* at elevated CO₂ concentrations. *Tree*

Physiology, 17(2), pp.89-96.

80. Arnone, J.A. and Gordon, J.C., 1990. Effect of nodulation, nitrogen fixation and CO₂ enrichment on the physiology, growth and dry mass allocation of seedlings of *Alnus rubra* Bong. *New Phytologist*, 116(1), pp.55-66.
81. Thomas, R.B., Bashkin, M.A. and Richter, D.D., 2000. Nitrogen inhibition of nodulation and N₂ fixation of a tropical N₂-fixing tree (*Gliricidia sepium*) grown in elevated atmospheric CO₂. *The New Phytologist*, 145(2), pp.233-243.
82. Quirk, J., Andrews, M.Y., Leake, J.R., Banwart, S.A. and Beerling, D.J., 2014. Ectomycorrhizal fungi and past high CO₂ atmospheres enhance mineral weathering through increased below-ground carbon-energy fluxes. *Biology letters*, 10(7), p.20140375.
83. Irigoyen, J.J., Goicoechea, N., Antolín, M.C., Pascual, I., Sánchez-Díaz, M., Aguirreolea, J. and Morales, F., 2014. Growth, photosynthetic acclimation and yield quality in legumes under climate change simulations: an updated survey. *Plant Science*, 226, pp.22-29.
84. Rogers, A., Ainsworth, E.A. and Leakey, A.D., 2009. Will elevated carbon dioxide concentration amplify the benefits of nitrogen fixation in legumes?. *Plant Physiology*, 151(3), pp.1009-1016.
85. Mohan, J.E., Clark, J.S. and Schlesinger, W.H., 2007. Long-term CO₂ enrichment of a forest ecosystem: implications for forest regeneration and succession. *Ecological Applications*, 17(4), pp.1198-1212.
86. Cernusak, L.A., Winter, K., Martínez, C., Correa, E., Aranda, J., Garcia, M., Jaramillo, C. and Turner, B.L., 2011. Responses of legume versus nonlegume tropical tree seedlings to elevated CO₂ concentration. *Plant physiology*, 157(1), pp.372-385.
87. Brantley, S.L., Megonigal, J.P., Scatena, F.N., Balogh-Brunstad, Z., Barnes, R.T., Bruns, M.A., Van Cappellen, P., Dontsova, K., Hartnett, H.E., Hartshorn, A.S. and Heimsath, A., 2011. Twelve testable hypotheses on the geobiology of weathering. *Geobiology*, 9(2), pp.140-165.
88. Vitousek, P.M. and Howarth, R.W., 1991. Nitrogen limitation on land and in the sea: how can it occur?. *Biogeochemistry*, 13(2), pp.87-115.
89. Barron, A.R., 2007. *Patterns and controls of nitrogen fixation in a lowland tropical forest, Panama*. Princeton University.
90. Wurzburger, N., Bellenger, J.P., Kraepiel, A.M. and Hedin, L.O., 2012. Molybdenum and phosphorus interact to constrain symbiotic nitrogen fixation in tropical forests. *PloS one*, 7(3), p.e33710.
91. López-Lara, I.M., Sohlenkamp, C. and Geiger, O., 2003. Membrane lipids in plant-associated bacteria: their biosyntheses and possible functions. *Molecular plant-microbe interactions*, 16(7), pp.567-579.
92. O'Hara, G.W., Dilworth, M.J., Boonkerd, N. and Parkpian, P., 1988. Iron-deficiency specifically limits nodule development in peanut inoculated with *Bradyrhizobium* sp. *New Phytologist*, 108(1), pp.51-57.
93. Gardner, W.K., Barber, D.A. and Parbery, D.G., 1983. The acquisition of phosphorus by *Lupinus albus* L. *Plant and soil*, 70(1), pp.107-124.
94. Roncal-Herrero, T. and Oelkers, E.H., 2011. Does variscite control phosphate availability in acidic natural waters? An experimental study of variscite dissolution rates. *Geochimica et Cosmochimica Acta*, 75(2), pp.416-426.
95. Wu, L., Huh, Y., Qin, J., Du, G. and van Der Lee, S., 2005. Chemical weathering in the Upper Huang He (Yellow River) draining the eastern Qinghai-Tibet Plateau. *Geochimica et Cosmochimica Acta*, 69(22), pp.5279-5294.

Figures (Epihov *et al.*, 2017)

Figure 1. Global rise of legume-rich tropical forests during the early Cenozoic (58 – 42 Myr ago). **a.** Global map of the major legume fossil records plotted on the Eocene continental configuration. Lines and their ball ends point to approximate locations. Caesalpinioids in the Wilcox flora are according the old pre-molecular taxonomy with a family status. Abbreviations: DTF = dry tropical forest, SubTF = sub-tropical forest, TRF = tropical rainforest, boreotropical or BTF = a forest with mixed tropical and temperate species which is sometimes referred to as boreotropical. **b.** Summary of the notable legume-rich fossil assemblages and all major molecular clock-dated crown nodes in the Leguminosae marking the rise of the legume-rich forests in the Palaeocene-Eocene plotted against atmospheric CO₂ records (light blue dots and red Loess curve) using data from [93] and ocean bottom water temperature (orange semi-transparent curve) using data from [94]. Peaks in terrestrial weathering (WTs = 55, 48, 35 Myr ago) are estimated as levels of lateritization and bauxitization in [30].

Abbreviations: Cjn = Cerrejon rainforest formation, Wlx = Wilcox boreotropical flora, Wy = Wyoming flora, Pat = Patagonia dry forests, Mah = Mahenge dry tropical forest, Cyn = *Cynometra*-monodominant stands in Mwadui, Cam = Cameroon tropical rainforest, Bjm = putative *Brachystegia-Julbernardia* miombo (macrofossils but not assemblage). Crown nodes include the divergence of L = Leguminosae, Pa = Papilionoideae, G = Genistoids, D = Dalbergioids, N = *Senna* clade, U = *Umtiza* clade, A = Amherstieae tribe (contains the majority EM taxa) after [95], S = *Swartzia* clade, R = Robinioids, B = Mirbelioids, I = Indigoferoids, Cl = *Cladrastis* clade, M = Millettoids, Mi = Mimosoideae, O = *Peltophorum* clade, T = *Trifolium* (IRLC) clade, C = *Cercis* clade, P = *Poeppigia* clade, F = Fossil-not-supported *Brachystegia* clade (because fossils of *Brachystegia* and *Julbernardia* found much earlier and new estimates show that this divergence occurred 52.1 Myr ago – here marked as clade Amherstieae). Clock data references: all clade ages unless otherwise stated are after [96].

Figure 2. Foliar N ratios between N₂-fixing and non-fixing non-legumes in (a) tropical forests and (b) tropical forestry plantations and between the three functional groups and (d) pathways of the nitrogen-weathering feedback hypothesis. Red typeface depicts

factors stimulating weathering with specific weathering reactions associated to those factors in brackets. In tropical forests, N₂-fixing legumes exhibit an average of 34.58% higher leaf crude protein content than non-fixing tree species. In forestry plantations, N₂-fixing legume species reveal on average 64.50% (SEM=11.57%) higher leaf crude protein content than non-fixing trees. Raw data and references are available in the Supplementary Information.

Figure 3. Atmospheric CO₂, NPP and weathering feedbacks. (a) Ecosystem effects of elevated CO₂ levels in legume-poor and rich forests; (b) differences in feedback relationships between rich and poor forests. In both forest types, high atmospheric CO₂ levels (1) promote a proportional NPP increase (2) which transitions the system to low N-availability (3). Ultimately, in poor forests that would result in a negative feedback on NPP. In rich forests, however, low N availability (3) can up-regulate N₂-fixation rates and recruitment of N₂-fixers (4) thus alleviating N limitations and allowing for unchanged CO₂-NPP relationship. Green arrows indicate positive relationships, whereas red ball-ending lines - negative relationships; N₂F = N₂-fixation.

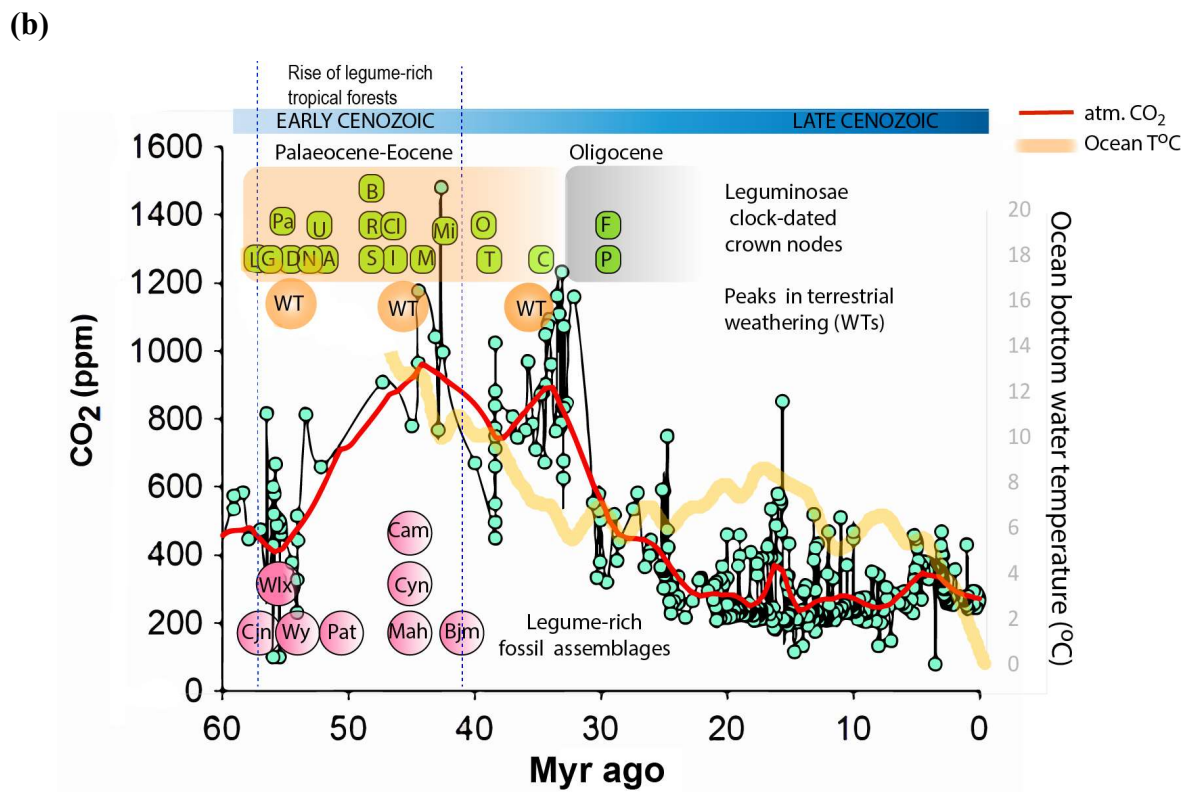
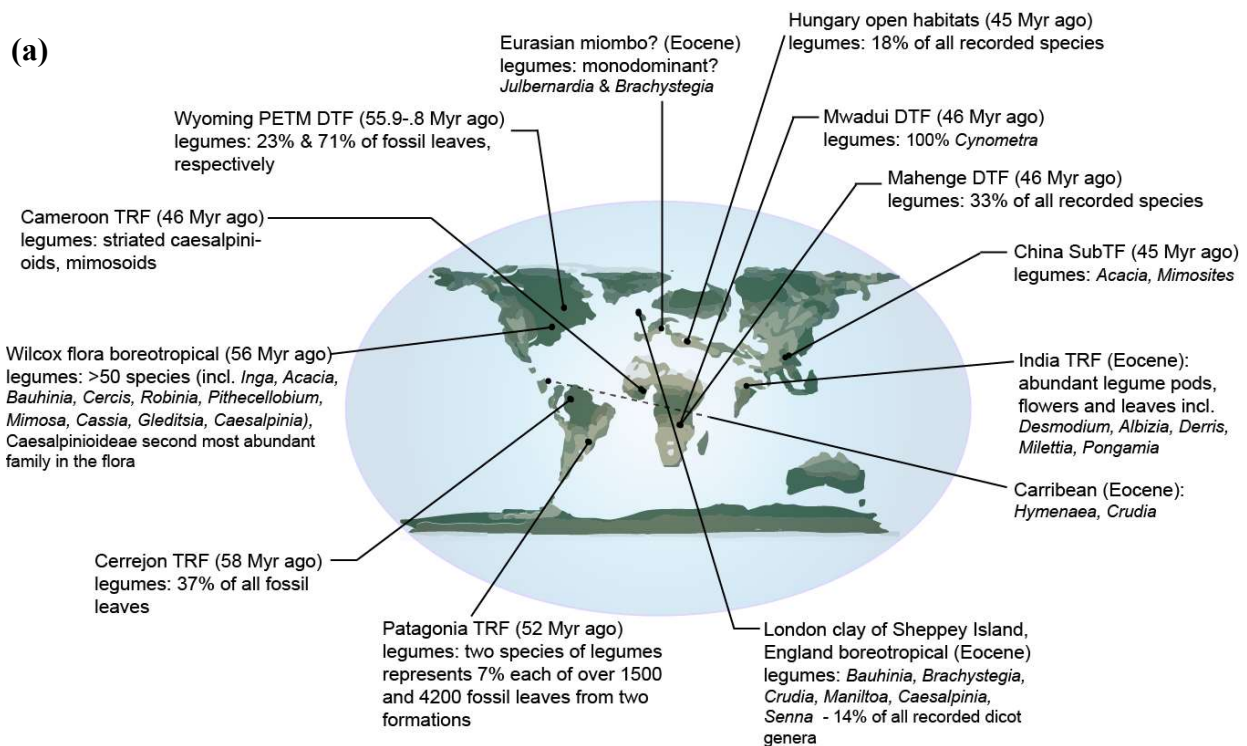


Figure 1
(Epihov *et al.*)

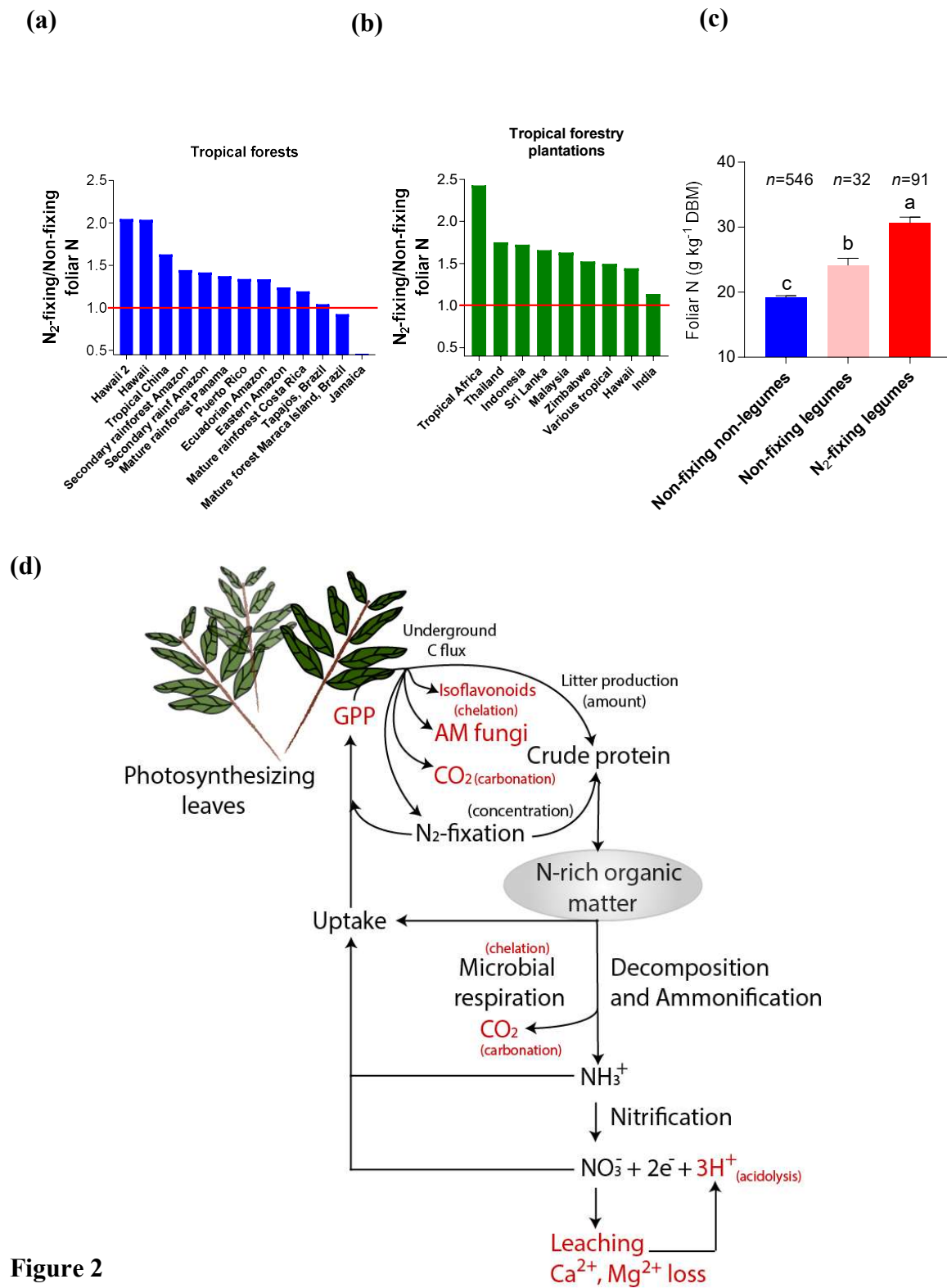


Figure 2
(Epihov *et al.*)

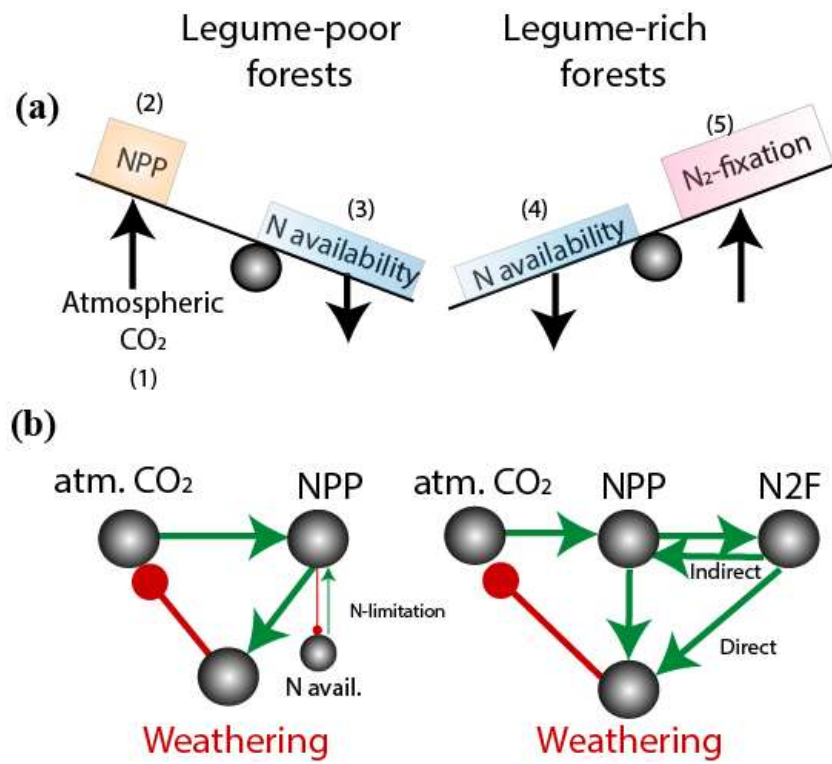


Figure 3
(Epihov *et al.*)

Supplementary Material for

N₂-fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?

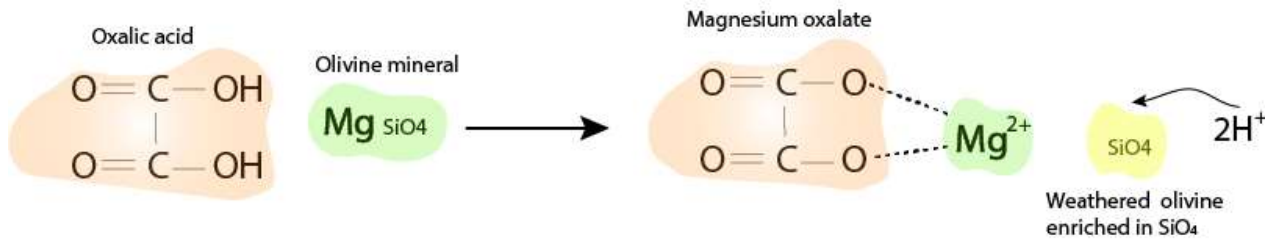
Dimitar Z. Epihov, Sarah A. Batterman, Lars O. Hedin, Jonathan R. Leake, Lisa M. Smith and David J. Beerling

Supplementary Figure 1. Illustrated glossary of weathering reactions

The following processes are often enhanced by biological (biotic) systems such as root exudation, microbial physiology, soil respiration and organic matter decomposition and leaching and therefore represent the basis of biological weathering [1].

Chelation (acido-complexolysis; complexation) – the reaction of complexing between metal ions from minerals/rocks with organic molecules (chelating agents; chelators) via the formation of coordination bonds. Important in biological weathering. Organic acids such as citric acid, oxalic acid, tartaric acid, acetic acid, lactic acid, gluconic acid and amino acids are major chelating agents. Chelating organic acids may also produce protons during their dissociation which can further attack minerals through acidolysis which is why some sources refer to chelation also as acido-complexolysis [2].

Suggested example: Oxalic acid + olivine \rightarrow $[\text{Mg}^{2+} : \text{oxalate}]$ complex + weathered olivine + 2H^+



Carbonation – CO_2 in soil produced by biological activity such as respiration often dissolves in water forming the weak carbonic acid (H_2CO_3). Carbonic acid can react with silicate minerals producing metal carbonates. The process can be sped-up by the presence of the enzyme carbonic anhydrase [3].

Suggested example:



Acidolysis (simple acidolysis, acid attack, protonation) – a process in which protons (generated biologically or by acid dissociation) replace the metal cations from mineral surfaces and bring the mineral metals to solution.

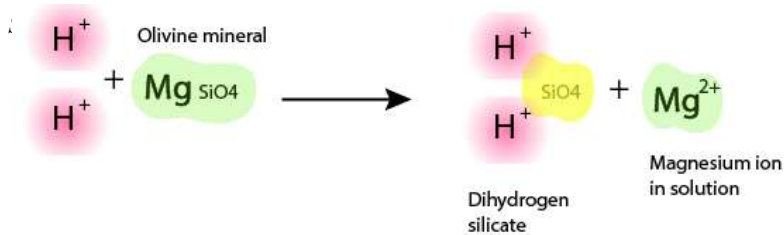
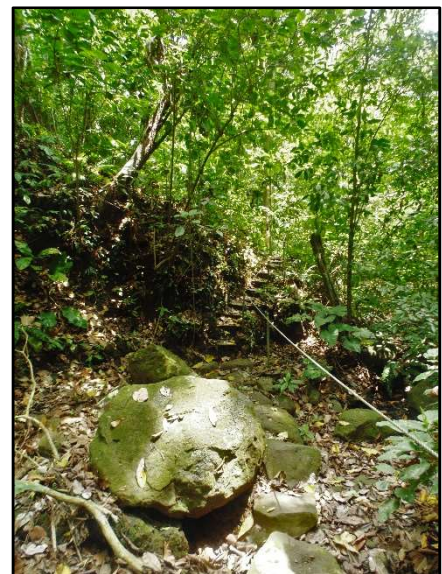


Figure I. Basalt boulders in the tropical forests of Barro Colorado Island, Panama – the presence of boulders (here seen with weathered surface of altered colour) may indicate that high amounts of volcanic mineral materials at different stages of weathering are contained within the soil horizons where chelation, acidolysis and carbonation driven by forest processes can stimulate their dissolution. Photo credit: Dimitar Z. Epihov.



Supplementary Figure 2. Functional symbiotic diversity within Leguminosae – coupling extant and fossil evidence

Root microbial symbioses can be divided into 2 major groups - dipartite (that is symbioses between a plant host and a single symbiont group) and multipartite (that is symbioses between a plant host and two or more symbiotic partners). Plants with dipartite symbioses include arbuscular mycorrhizal (AM) plants, ectomycorrhizal (EM), ericoid mycorrhizal (ERM) etc. Plants with multipartite symbioses are the group of N₂-fixing and arbuscular mycorrhizal plants (NAM), N₂-fixing and ectomycorrhizal plants (NEM), and N₂-fixing dual arbuscular-ectomycorrhizal plants (NAEM).

Leguminosae is one of the most symbiotically-rich plant families with members known to exhibit AM, EM, NAM, NEM [4],[5] or NAEM [6] properties. Symbiotic assignment to fossil taxa is based upon the symbiotic characteristics of extant members of that taxon as found in the Supplementary curated by Werner *et al.* in [7] (except *Xylia* - [8] and *Maniltoa* - [9]). Lists of ectomycorrhizal legumes are found in [10] and [11].

Functional type

NAM

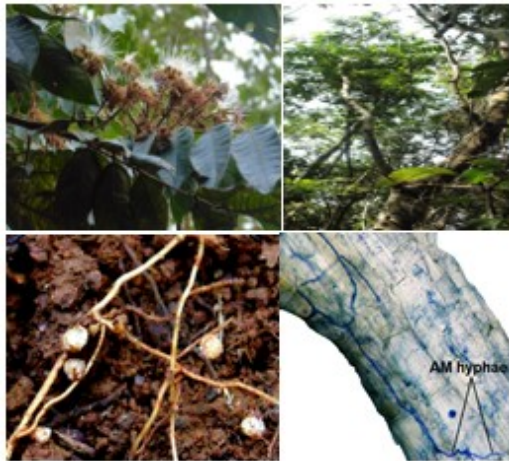


Figure 1. N₂-fixing AM legumes of Neotropical rainforests: *Inga cocleensis* – an extant member of the ancient *Inga* genus with origins in the early Cenozoic. Shown here are flowers and leaves, canopy, *Burkholderia* root nodules and Trypan Blue-stained intraradical AM hyphae. Photo credit: Dimitar Z. Epihov

Palaeocene/Eocene Fossil Legume Taxa

Inga [12],[13],[14], *Prosopis* [15],[16], *Mimosa* [15], *Chamaechrista* [12], *Acacia* [13],[16],[17],[18],[19], *Swartzia* [16],[20], *Albizia* [15],[16], *Pentaclethra* [15],[19], *Adentanthera* [15],[19] *Ormosia* [16], *Sophora* [13],[16], *Robinia* [16],[20], *Diploptropis* [16], *Canavalia* [13], *Dalbergia* [13],[20], *Machaerium* [16], *Strongyolobos* [21], *Pongamia* [21], *Neptunia* [16], *Derris* [22], *Desmodium* [22], *Milletia* [22], *Maniltoa* [23], *Crudia* [23],[24], *Xylia* [16]

EM

Aphanocalyx [17], *Afzelioxylon*/*Afzelia* [25],[26], *Brachystegia* [27],[28], *Julbernardia* [29]

AM

Ablygonocarpus [15], *Gymnocladus* [20], *Cladrastis* [20], *Senna* [16], *Calpocalyx* [15], *Cassia* [30], *Cynometra* [31], *Peltogyne* [16], *Bauhinia* [20], *Vouapa* [21], *Hymenaea* [23], *Caesalpinia* [16]

Modern analogues of Cenozoic fossil forests. The occurrence of fossils of the above taxa at different sites suggests early Cenozoic forests exhibited compositional patterns analogous to major types of modern forests with Leguminosae as an important family in both species-rich NAM forests and monodominant forests of EM or AM legumes. We suggest fossils of *Inga*, *Swartzia*, *Machaerium* and the AM legume taxa (*Senna*, *Cassia*) might be analogous to species-rich NAM legume tropical rainforests of modern Amazon, Panama and Costa Rica [32]. Fossils of *Brachystegia* and *Julbernardia* are suggestive of monodominant EM dry tropical forests like miombo woodlands in Africa [33]. Abundant *Acacia* fossil records are often interpreted as dry tropical forests analogous to savanna *Acacia* woodland communities currently found in Africa and Mexico [34]. Fossil assemblages of *Cynometra*- and pollen of *Peltogyne*-affinity might be analogous to the monodominant legume AM communities forming *Cynometra alexandrii* forests in Africa and that of *Peltogyne gracilipes* in Amazon [35]. Finally, we suggest that the presence of fossil EM *Aphanocalyx* and *Afzelia* might be indicative as accessory EM species found in modern monodominant EM rainforests like the EM legume *Microberlinia bisulcata* monodominant forests in Cameroon [11],[36].

Supplementary References

1. Taylor, L.L., Leake, J.R., Quirk, J., Hardy, K., Banwart, S.A. and Beerling, D.J., 2009. Biological weathering and the long-term carbon cycle: integrating mycorrhizal evolution and function into the current paradigm. *Geobiology*, 7(2), pp.171-191.
2. Van Rompaey, K., Van Ranst, E., Verdoodt, A. and De Coninck, F., 2007. Use of the test-mineral technique to distinguish simple acidolysis from acido-complexolysis in a Podzol profile. *Geoderma*, 137(3-4), pp.293-299.
3. Li, W., Yu, L.J., He, Q.F., Wu, Y., Yuan, D.X. and Cao, J.H., 2005. Effects of microbes and their carbonic anhydrase on Ca^{2+} and Mg^{2+} migration in column-built leached soil-limestone karst systems. *Applied Soil Ecology*, 29(3), pp.274-281.
4. Hogberg, P., 1986. Nitrogen-fixation and nutrient relations in savanna woodland trees (Tanzania). *Journal of Applied Ecology*, pp.675-688.
5. Diédhiou, A.G., Guèye, O., Diabaté, M., Prin, Y., Duponnois, R., Dreyfus, B. and Bâ, A.M., 2005. Contrasting responses to ectomycorrhizal inoculation in seedlings of six tropical African tree species. *Mycorrhiza*, 16(1), pp.11-17.
6. Hopkins, M.S., Reddell, P., Hewett, R.K. and Graham, A.W., 1996. Comparison of root and mycorrhizal characteristics in primary and secondary rainforest on a metamorphic soil in North Queensland, Australia. *Journal of Tropical Ecology*, 12(6), pp.871-885.
7. Werner, G.D., Cornwell, W.K., Sprent, J.I., Kattge, J. and Kiers, E.T., 2014. A single evolutionary innovation drives the deep evolution of symbiotic N_2 -fixation in angiosperms. *Nature Communications*, 5, p.4087.
8. Teamroong, N. and Boonkerd, N., 2006. Rhizobial production technology. In *Microbial Biotechnology in Agriculture and Aquaculture*, Vol. 2 (pp. 89-122). CRC Press.
9. Lewin, A., Rosenberg, C., Wong, C.H., Nelson, L., Manen, J.F., Stanley, J., Dowling, D.N., Dénarié, J. and Broughton, W.J., 1987. Multiple host-specificity loci of the broad host-range *Rhizobium* sp. NGR234 selected using the widely compatible legume *Vigna unguiculata*. *Plant molecular biology*, 8(6), pp.447-459.
10. Alexander, I.J. and Högborg, P., 1986. Ectomycorrhizas of tropical angiospermous trees. *New Phytologist*, 102(4), pp.541-549.
11. Newbery, D.M., Alexander, I.J., Thomas, D.W. and Gartlan, J.S., 1988. Ectomycorrhizal rain-forest legumes and soil phosphorus in Korup National Park, Cameroon. *New Phytologist*, 109(4), pp.433-450.
12. Calvillo-Canadell, L. and Cevallos-Ferriz, S.R.S., 2005. Diverse assemblage of Eocene and Oligocene leguminosae from Mexico. *International Journal of Plant Sciences*, 166(4), pp.671-692.
13. Berry, E.W., 1930. *Revision of the lower Eocene Wilcox flora of the southeastern states: with descriptions of new species, chiefly from Tennessee and Kentucky* (Vol. 156). US Government Printing Office.
14. Irving, R.S. and Stuessy, T.F., 1971. A New Paratropical Angiosperm Florule in the Eocene Rockdale Formation of Bastrop County, Texas. *The Southwestern Naturalist*, pp.111-116.
15. Caccavari, M.A., 1996. Analysis of the South American fossil pollen record of Mimosoideae (Leguminosae). *Review of Palaeobotany and Palynology*, 94(1-2), pp.123-135.
16. Herendeen, P.S., 1992. The fossil history of the Leguminosae: phylogenetic and biogeographic implications. In *Advances in Legume Systematics, Part 4. The Fossil Record* (pp. 303-316). Royal Botanic Gardens.
17. Herendeen, P.S. and Jacobs, B.F., 2000. Fossil legumes from the middle Eocene (46.0 Ma) Mahenge flora of Singida, Tanzania. *American Journal of Botany*, 87(9), pp.1358-1366.
18. Wang, Q., Ferguson, D.K., Feng, G.P., Ablav, A.G., Wang, Y.F., Yang, J., Li, Y.L. and Li, C.S., 2010. Climatic change during the Palaeocene to Eocene based on fossil plants from Fushun, China. *Palaeogeography, Palaeoclimatology, Palaeoecology*, 295(1-2), pp.323-331.
19. Maley, J., 1996. The African rain forest—main characteristics of changes in vegetation and climate from the Upper Cretaceous to the Quaternary. *Proceedings of the Royal Society of Edinburgh, Section B: Biological Sciences*, 104, pp.31-73.
20. Taylor, D.W., 1990. Paleobiogeographic relationships of angiosperms from the Cretaceous and early Tertiary of the North American area. *The Botanical Review*, 56(4), pp.279-417.
21. MacGinitie, H.D., 1941. Middle Eocene flora from the central Sierra Nevada.
22. Awasthi, N., 1992. Indian fossil legumes. *Advances in legume Systematics*, 225, p.250.
23. Graham, A.L.A.N., 1992. The current status of the legume fossil record in the Caribbean region. *Advances in legume systematics, part, 4*, pp.161-167.
24. Herendeen, P.S., 1992. Early caesalpinoid fruits from the Palaeogene of southern England. In *Advances in Legume Systematics, part 4. The Fossil Record* (pp. 57-68). Royal Botanic Gardens.
25. Pan, A.D., Jacobs, B.F. and Herendeen, P.S., 2010. Detarieae *sensu lato* (Fabaceae) from the Late

CHANGED!

- Oligocene (27.23 Ma) Guang River flora of north-western Ethiopia. *Botanical Journal of the Linnean Society*, 163(1), pp.44-54..
26. In press. *Afzelioxylon furoni* (N° inventaire 5731)
<http://albinoni.snv.jussieu.fr/DB/specimens/fiche/spe442.html>.
 27. Shakryl, A.K., 1992. Leguminosae species from the Tertiary of Abkhazia. *Advances in legume systematics, part, 4*, pp.189-206.
 28. Ettingshausen, C.B., 1879. III. Report on Phyto-Palæontological investigations of the fossil Flora of Sheppey. *Proceedings of the Royal Society of London*, 29(196-199), pp.388-396.
 29. Mijarra, J.M.P., Barrón, E., Manzanque, F.G. and Morla, C., 2009. Floristic changes in the Iberian Peninsula and Balearic Islands (south-west Europe) during the Cenozoic. *Journal of Biogeography*, 36(11), pp.2025-2043.
 30. Wilf, P., Johnson, K.R., Cuneo, N.R., Smith, M.E., Singer, B.S. and Gandolfo, M.A., 2005. Eocene plant diversity at Laguna del Hunco and Río Pichileufú, Patagonia, Argentina. *The American Naturalist*, 165(6), pp.634-650.
 31. Cantrill, D.J., Bamford, M.K., Wagstaff, B.E. and Sauquet, H., 2013. Early Eocene fossil plants from the Mwadui kimberlite pipe, Tanzania. *Review of palaeobotany and palynology*, 196, pp.19-35.
 32. Batterman, S.A., 2013. *Symbiotic nitrogen fixation in tropical forests: Scaling from individuals to ecosystems* (Doctoral dissertation, Princeton University).
 33. Malmer, A., 2007. General ecological features of miombo woodlands and considerations for utilization and management. In *MITMIOMBO—management of indigenous tree species for ecosystem restoration and wood production in semi-arid miombo woodlands in Eastern Africa. Proceedings of the First MITMIOMBO Project Workshop. Morogoro, Tanzania* (pp. 34-42).
 34. Greenberg, R., Bichier, P. and Sterling, J., 1997. *Acacia*, cattle and migratory birds in southeastern Mexico. *Biological Conservation*, 80(3), pp.235-247.
 35. Torti, S.D. and Coley, P.D., 1999. Tropical Monodominance: A Preliminary Test of the Ectomycorrhizal Hypothesis 1. *Biotropica*, 31(2), pp.220-228.
 36. Bechem, E.E.T., Chuyong, G.B. and Fon, B.T., 2014. A survey of mycorrhizal colonization in the 50-ha Korup Forest Dynamic Plot in Cameroon. *American Journal of Plant Sciences*, 5(10), p.1403.

Chapter 1

Legume-microbial interactions unlock mineral nutrients during tropical forest succession

Dimitar Z. Epihov^{1*}, Sarah A. Batterman², Kristin Saltonstall³, Lars O. Hedin⁴, Jefferson S. Hall³, Jonathan R. Leake¹, and David J. Beerling¹

¹*Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK*

²*School of Geography, University of Leeds, Leeds LS2 9JT, UK*

³*Smithsonian Tropical Research Institute, Balboa, Ancón, Panamá, Panama*

⁴*Department of Ecology and Evolutionary Biology, Princeton University, Princeton, New Jersey 08544, USA*

*for correspondence

Legume trees form an abundant and functionally important component of tropical forests worldwide^{1,2,3} with many developing N₂-fixing symbioses that promote higher growth and recruitment rates than non-fixers during early secondary succession^{4,5}. However, it remains unclear how N₂-fixers meet the elevated demands for inorganic nutrients that result from high rates of growth and biomass accumulation. Here we show that N₂-fixing trees in secondary Neotropical forests have 2-fold higher weathering rates of *in situ* fresh silicate minerals than non-fixers. Sequencing and annotation of 12 full shotgun metagenomes from weathered minerals indicate distinctive mineral microbiomes of N-fixers that support the hypothesized roles⁶ of enhanced nitrogen and carbon cycling (carbohydrate metabolic and respiratory potential) in weathering by generating localized acidity. Mineral-associated metagenomes and microbial communities of soils beneath N₂-fixers were linked to reductive conditions and fermentative acid products of fast-decomposing N-rich litter that favour phosphorus dissolution from minerals found in tropical soils. Furthermore, our analyses indicate the transfer of weathering benefits of N₂-fixers to non-fixing neighbouring trees and therefore the wider biogeochemical functioning of tropical forest ecosystems.

The legume family is the most numerous angiosperm family in Neotropics^{1,2}. Successful deployment of the N₂-fixing strategy of legumes in tropical forests^{2,3,4} depends on access to scarce inorganic nutrient sources, including bioavailable forms of phosphorus (P) and molybdenum (Mo)⁷, the latter being essential for synthesis of nitrogenases⁸. However, P and

Mo are often occluded in highly insoluble iron (Fe) and aluminium (Al)-bearing minerals^{9,10} and thus not available for immediate biological uptake. This situation is exemplified by the secondary Neotropical moist forests in Panama, where soils are P-poor oxisols, which developed on Mo-poor basalt bedrock containing large amounts of P-adsorbing kaolinite and goethite secondary minerals (**Supplementary Note 1**). Soil beneath N₂-fixers in the 17-year old Panamanian secondary forests had significantly lower pH, total P and Fe than those beneath non-fixers (**Supplementary Table 2**). Total soil P in these forests strongly correlated with total soil Fe (mainly goethite) and Al (mainly kaolinite) but not total soil organic C (Pearson test, $P < 0.001$ for Fe and Al and $P > 0.10$ for C; **Supplementary Figure 3**), suggesting that decline in total P in these forests is consistent with decline in Al and Fe-bearing minerals. These patterns raise the biogeochemical hypothesis that N₂-fixing legume trees may employ enhanced mineral weathering to access occluded inorganic mineral nutrients^{6,11} and sequester carbon into biomass during secondary succession.

Here we address this hypothesis by investigating (1) whether N₂-fixers exert higher silicate mineral weathering rates than non-fixing trees, with depletion of elements important for N₂ fixation, (2) whether enhanced weathering rates and N₂-fixing status are linked to functional differences in the structure and function of mineral microbiomes and metagenomes and (3) how N₂-fixers affect the biogeochemical potential of soil microbiota. We investigated these questions by undertaking a field study in Panama with replicated experimental plots (0.1 ha each, $n = 6$) representing a natural gradient in N₂-fixer abundance (6-27% tree basal area) in secondary Neotropical moist forests. This design allowed us to test whether N₂-fixing trees influence the soil microbiome and weathering rates in a manner that favours their rapid growth, and also exert community effects on weathering rates and soil microbiomes of nearby non-fixing trees relative to non-fixers far from fixers in legume-poor forests. We determined *in situ* weathering rates of trees in Panama by burying >500 mesh bags containing crushed dunite, an olivine-rich (>90%) silicate rock (**Supplementary Table 2**) in the rooting zone (~10 cm depth) of N₂-fixing (5 species, $n = 51$ trees), non-fixing trees far from N₂-fixers in legume-poor forests (NF-far) (5 species, $n = 36$ trees) and non-fixing trees near fixers in legume-rich forests (NF-near) (5 species, $n = 39$ trees). After 8 months, recovered samples ($n = 126$ samples) were analysed to calculate olivine weathering rates relative to fresh unweathered olivine samples.

X-ray fluorescence (XRF) revealed that the soil beneath N₂-fixing legumes had double the olivine weathering rate than soil from non-fixers (Welch's t-test, $n=126$, $P=0.026$;

Supplementary Note 4) (**Figure 1a**). Community effects of N₂-fixing trees on weathering rates were also apparent, with non-fixing trees near N₂-fixing trees having weathering rates intermediate between N₂-fixers and non-fixing trees located far from fixers (ANOVA $P < 0.05$, Fisher's LSD test) (**Figure 1b**). High olivine weathering rates were associated to significant decrease in mineral Mo% content (Spearman test, $P < 0.05$, **Supplementary Figure 5**) suggesting depletion of elements important for N₂-fixation. Overall, olivine weathering rates were linked to declines in the pH of the soil (Spearman test, $P < 0.01$; **Supplementary Figure 5**) and reacted olivine, notably with conditions beneath N₂-fixing trees being significantly more acidic than beneath non-fixing trees, in-line with symbiotic nitrogen fixation^{6,12,13} (two-tailed t-test, $P < 0.05$; **Figure 1c**). Further production of N₂-fixing root nodules (hereafter "nodulation") was a significant factor affecting the nickel (Ni)% content of olivine (two-tailed t-test, $P < 0.05$ and ANOVA, $P < 0.05$ in **Figure 1e-f**, respectively). Increasing Ni% content of olivine is a hallmark of intensifying weathering converting olivine material into secondary nickiliferous clay minerals¹⁴. Olivine Ni% content indeed negatively correlated with weathering rates in our experiment (Spearman test, $P < 0.01$; **Supplementary Figure 5**) suggesting that sites of active nodulation promote regimes of enhanced weathering in agreement to their H⁺-exuding activities¹⁵. Across the natural gradient in N₂-fixer abundance in our field experiment, soil C:N, a parameter reflective of high N inputs¹⁶, was lower beneath N₂-fixers and NF-near compared to NF-far, indicating a second community effect in addition to soil pH (**Figure 1d**). Soil C:N ratios also correlated with weathering rates (Spearman test, $P < 0.01$; **Supplementary Figure 5**) highlighting the importance of C and N cycling to weathering.

To analyse the role of the microbiome in mineral dissolution and establish the biogeochemical conditions favouring weathering, we constructed and sequenced 12 shotgun metagenome libraries of reacted olivine mineral samples from beneath N₂-fixers ($n = 6$), NF-far ($n = 3$) and NF-near ($n = 3$). Metagenomes provided information on the abundance of gene orthologues representing 195 functional metabolic pathways (Level 2 in Subsystem MG-RAST database¹⁷) that were combined with weathering rates and analysed using hierarchical clustering to objectively assess which pathways were associated with weathering, as expressed in a correlation heat-map (**Figure 2a**). Results showed a single specific well-supported cluster linking enhanced weathering rates with coordinated increases in gene abundance for pathways involved in microbial respiration, carbohydrate metabolism and N cycling (**Figure 2a**). Analysis of these metagenomes normalized using single copy marker

gene (*rpoC*)¹⁸ reinforced this view, identifying significant correlations between *in situ* olivine weathering beneath N₂-fixers, NF-far and NF-near and gene abundance representing lower level metabolic pathways for N and C cycling (**Figure 2b**) (for single gene correlations refer to Supplementary Table 7).

Detailed analysis of the microbial N cycling pathways indicated the abundance of genes involved in nitrification, assimilatory and dissimilatory N reduction, denitrification and ammonia assimilation all correlate with weathering (**Figures 2b,c**). Mechanistically, this metabolic profile can be understood in terms of increased inputs of fixed N beneath N₂-fixers stimulating N cycling and H⁺ generation promoted by nitrate leaching^{6,12,13} (**Figure 2c**). For microbial C cycling, significant positive correlations occurred between weathering and gene abundance linked to (1) C storage pathways (glycogen and starch synthesis), (2) Krebs cycle enzyme-coding genes, (3) genes of the Entner-Doudoroff-type of glucose breakdown, (4) sugar-transporting machinery genes (**Figure 2b, 2d**; Supplementary Table 7) that are all indicative of high C availability^{19,20}. Likely in response to high C flux to microbes, gene abundance in the microbial respiration (electron-transport chain) and Krebs cycle, pathways understood to drive respiratory CO₂ fluxes and generate acidity, also correlated with weathering (**Figure 2a,b,d**). The tight coupling between N, C cycling, the availability of those two macroelements and weathering is revealed by genes linking the interconversion of amino acids to organic acid intermediates through the Krebs cycle (thus controlling N-C feedbacks on bacterial metabolism²¹), in particular their positive association with weathering rates (**Figure 2e**). Overall, patterns of gene abundance relating to microbial N cycling, respiration and the Krebs cycle follow consistent trends of N₂-fixers > NF-near > NF-far (**Figure 2c**) which supports N₂-fixer-driven community effects on weathering rates (**Figure 1b**).

A characteristic feature of the low C:N ratio of legume litter and its resulting soil organic matter⁶ (**Figure 1d**) is rapid decomposition and microbial activity^{22,23} that can result in localized depletion of available O₂, formation of anaerobic microsites and anaerobiosis in soil and mineral aggregates²⁴. Anaerobic conditions can lead to Fe reduction, generation of fermentative acid products and S cycle H₂S corrosive species that can all contribute to weathering, and particularly the dissociation of P from its adsorbed sites on Fe oxides, particularly relevant in this P-limited soil matrix. We therefore investigate interactions between weathering and gene abundance for major anaerobic processes and discovered correlations with marker genes for the microbial cycling of Fe and S as well as mixed acid fermentation (**Figure 2b**). Moreover, our metagenomic analyses showed positive correlations

between the relative abundance of genes coding for Mo-containing enzymes and olivine weathering (**Supplementary Figure 6**), indicating that weathering supplies Mo for the necessary co-factors to synthesize anaerobic enzymes²⁵. Enhanced dissolution of Mo from minerals may therefore stimulate symbiotic N₂-fixation and augment mineral weathering via positive feedbacks on anaerobic metabolic processes.

To gain insights into microbial weathering processes allowing trees to access P occluded in insoluble Fe and Al-bearing minerals not present in olivine, we undertook determinations of the microbiome structure and function in soils beneath our three categories of trees (N₂-fixers, NF-far, NF-near; *n* = 21, 12 and 13 samples, respectively), using Next-Generation 16S rRNA sequencing. Microbial profiling data show that for N and S cycling lineages, N₂-fixers exhibit significant community effects on near-by neighbouring trees (NF-near) but not on NF-far trees (**Figure 3a,b**). N-cycling microbial communities in NF-near and N₂-fixers were enriched in ammonia oxidisers, particularly the archaeal genus of *Nitrososphaera*, which converts NH₃ to NO₂⁻ and generates acidity^{26,27}. S-cycling communities were dominated by *Thioalkalivibrio* which oxidizes S to SO₄²⁻ and also generates acidity (two protons per reaction)²⁸. **Fe cycling microorganisms appear to exhibit higher cumulative abundance in soils beneath N₂-fixers and NF-near relative to NF-far, once again revealing apparent community effects** (**Figure 3c**). Changes in the microbial communities involved in Fe and S cycling can affect Fe-bound P by reducing the highly insoluble Fe(III)P to Fe (II) + P (Fe cycling)²⁹ and Fe(II)S + P (S cycling)^{30,31}. Together, these findings indicate the rhizospheric soil microbiome of N₂-fixers improves access to highly limiting P, a benefit that may also be shared with neighbouring non-fixers.

We conclude that the functional soil microbial community of N₂ fixing trees provides them with access to mineral resources to help meet their high demands for fast growth. Metagenomics indicates faster weathering of N₂-fixing trees arises through upregulation of specific classes of metabolic pathways linking microbial energy metabolism with inorganic mineral nutrient cycling. These lead to (1) acidification of the immediate matrix by CO₂ evolution and carbonic acid production, (2) the formation of anaerobic microsites conducive to reductive dissolution resulting from Fe and S cycling affecting the mineral lattice, (3) the acid attack by fermentative and Krebs cycle acid products and (4) generation of excess H⁺ by enhanced NO₃⁻ leaching as consequence of greater N inputs and nitrification. The formation of anaerobic niches appears under the control of Mo supply required for nodulation in legume trees and mineral weathering. Further, we have shown community effects of N₂-fixers on

weathering rates, soil physicochemical factors filtered down to community effects in metagenomes and soil microbiomes, highlighting previously overlooked roles (with most previous studies mainly focusing on rhizobia-legume interactions in microbiomes^{32,33}) of these trees in tropical forest nutrient cycling and biogeochemistry.

References

1. Ter Steege, H., Pitman, N.C., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.F., Prévost, M.F., Spichiger, R., Castellanos, H. and Von Hildebrand, P., 2006. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, 443(7110), p.444.
2. Hedin, L.O., Brookshire, E.J., Menge, D.N. and Barron, A.R., 2009. The nitrogen paradox in tropical forest ecosystems. *Annual Review of Ecology, Evolution, and Systematics*, 40, pp.613-635.
3. Yeo, W.L.J. and Fensham, R.J., 2014. Will Acacia secondary forest become rainforest in the Australian Wet Tropics?. *Forest ecology and management*, 331, pp.208-217.
4. Batterman, S.A., Hedin, L.O., Van Breugel, M., Ransijn, J., Craven, D.J. and Hall, J.S., 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature*, 502(7470), p.224.
5. Menge, D.N. and Chazdon, R.L., 2016. Higher survival drives the success of nitrogen-fixing trees through succession in Costa Rican rainforests. *New Phytologist*, 209(3), pp.965-977.
6. Epihov, D.Z., Batterman, S.A., Hedin, L.O., Leake, J.R., Smith, L.M. and Beerling, D.J., 2017. N₂-fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?. *Proceedings of the Royal Society B: Biological Sciences*, 284(1860), p.20170370.
7. Wurzburger, N. and Hedin, L.O., 2016. Taxonomic identity determines N₂ fixation by canopy trees across lowland tropical forests. *Ecology letters*, 19(1), pp.62-70.
8. Barron, A.R., Wurzburger, N., Bellenger, J.P., Wright, S.J., Kraepiel, A.M. and Hedin, L.O., 2009. Molybdenum limitation of symbiotic nitrogen fixation in tropical forest soils. *Nature Geoscience*, 2(1), p.42.
9. Chacon, N., Silver, W.L., Dubinsky, E.A. and Cusack, D.F., 2006. Iron reduction and soil phosphorus solubilization in humid tropical forests soils: the roles of labile carbon pools and an electron shuttle compound. *Biogeochemistry*, 78(1), pp.67-84.
10. Goldberg, S., Forster, H.S. and Godfrey, C.L., 1996. Molybdenum adsorption on oxides, clay minerals, and soils. *Soil Science Society of America Journal*, 60(2),

pp.425-432.

11. Nasto, M.K., Alvarez-Clare, S., Lekberg, Y., Sullivan, B.W., Townsend, A.R. and Cleveland, C.C., 2014. Interactions among nitrogen fixation and soil phosphorus acquisition strategies in lowland tropical rain forests. *Ecology letters*, 17(10), pp.1282-1289.
12. Tang, C., Unkovich, M.J. and Bowden, J.W., 1999. Factors affecting soil acidification under legumes. III. Acid production by N₂-fixing legumes as influenced by nitrate supply. *The New Phytologist*, 143(3), pp.513-521.
13. Van Miegroet, H. and Cole, D.W., 1984. The Impact of Nitrification on Soil Acidification and Cation Leaching in a Red Alder Ecosystem 1. *Journal of Environmental Quality*, 13(4), pp.586-590.
14. Nahon, D.B., Paquet, H. and Delvigne, J., 1982. Lateritic weathering of ultramafic rocks and the concentration of nickel in the western Ivory Coast. *Economic Geology*, 77(5), pp.1159-1175.
15. Ding, X., Sui, X., Wang, F., Gao, J., He, X., Zhang, F., Yang, J. and Feng, G., 2012. Synergistic interactions between *Glomus mosseae* and *Bradyrhizobium japonicum* in enhancing proton release from nodules and hyphae. *Mycorrhiza*, 22(1), pp.51-58.
16. Cools, N., Vesterdal, L., De Vos, B., Vanguelova, E. and Hansen, K., 2014. Tree species is the major factor explaining C:N ratios in European forest soils. *Forest Ecology and Management*, 311, pp.3-16.
17. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A. and Wilkening, J., 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9(1), p.386.
18. Aserse, A. A., Räsänen, L. A., Aseffa, F., Hailemariam, A. & Lindström, K. Diversity of sporadic symbionts and nonsymbiotic endophytic bacteria isolated from nodules of woody, shrub, and food legumes in Ethiopia. *Appl. Microbiol. Biotechnol.* **97**, 10117–10134 (2013).
19. Aserse, A.A., Räsänen, L.A., Aseffa, F., Hailemariam, A. and Lindström, K., 2013. Diversity of sporadic symbionts and nonsymbiotic endophytic bacteria isolated from nodules of woody, shrub, and food legumes in Ethiopia. *Applied microbiology and biotechnology*, 97(23), pp.10117-10134.
20. Preiss, J., 1984. Bacterial glycogen synthesis and its regulation. *Annual review of microbiology*, 38(1), pp.419-458.
21. Shimizu, K., 2014. Regulation systems of bacteria such as *Escherichia coli* in response to nutrient limitation and environmental stresses. *Metabolites*, 4(1), pp.1-35.
22. Milcu, A., Partsch, S., Scherber, C., Weisser, W.W. and Scheu, S., 2008. Earthworms and legumes control litter decomposition in a plant diversity gradient. *Ecology*, 89(7), pp.1872-1882.
23. Schwendener, C.M., Lehmann, J., Rondon, M., Wandelli, E. and Fernandes, E., 2007.

- Soil mineral N dynamics beneath mixtures of leaves from legume and fruit trees in Central Amazonian multi-strata agroforests. *Acta Amazonica*, 37(3), pp.313-320.
24. Young, I.M. and Crawford, J.W., 2004. Interactions and self-organization in the soil-microbe complex. *Science*, 304(5677), pp.1634-1637.
 25. Hover, B.M., Tonthat, N.K., Schumacher, M.A. and Yokoyama, K., 2015. Mechanism of pyranopterin ring formation in molybdenum cofactor biosynthesis. *Proceedings of the National Academy of Sciences*, 112(20), pp.6347-6352.
 26. Tournai, M., Stieglmeier, M., Spang, A., Könneke, M., Schintlmeister, A., Urich, T., Engel, M., Schlöter, M., Wagner, M., Richter, A. and Schleper, C., 2011. *Nitrososphaera viennensis*, an ammonia oxidizing archaeon from soil. *Proceedings of the National Academy of Sciences*, 108(20), pp.8420-8425.
 27. Szwerinski, H., Arvin, E. and Harremoës, P., 1986. pH-decrease in nitrifying biofilms. *Water Research*, 20(8), pp.971-976.
 28. Sorokin, D.Y., Tourova, T.P., Lysenko, A.M., Mityushina, L.L. and Kuenen, J.G., 2002. *Thioalkalivibrio thiocyanoxidans* sp. nov. and *Thioalkalivibrio paradoxus* sp. nov., novel alkaliphilic, obligately autotrophic, sulfur-oxidizing bacteria capable of growth on thiocyanate, from soda lakes. *International journal of systematic and evolutionary microbiology*, 52(2), pp.657-664.
 29. Chacon, N., Silver, W.L., Dubinsky, E.A. and Cusack, D.F., 2006. Iron reduction and soil phosphorus solubilization in humid tropical forests soils: the roles of labile carbon pools and an electron shuttle compound. *Biogeochemistry*, 78(1), pp.67-84.
 30. Chi, R.A., Xiao, C.Q. and Gao, H., 2006. Bioleaching of phosphorus from rock phosphate containing pyrites by *Acidithiobacillus ferrooxidans*. *Minerals Engineering*, 19(9), pp.979-981.
 31. Hamilton, W.A., 1985. Sulphate-reducing bacteria and anaerobic corrosion. *Annual review of microbiology*, 39(1), pp.195-217.
 32. Wolińska, A., Kuźniar, A., Zielenkiewicz, U., Banach, A., Izak, D., Stępniewska, Z. and Błaszczuk, M., 2017. Metagenomic analysis of some potential nitrogen-fixing bacteria in arable soils at different formation processes. *Microbial ecology*, 73(1), pp.162-176.
 33. Dinnage, R., Simonsen, A.K., Cardillo, M., Thrall, P.H., Barrett, L.G. and Prober, S.M., 2018. Larger legume plants host a greater diversity of symbiotic nitrogen-fixing bacteria. *bioRxiv*, p.246611.
 34. Poorter, L., Bongers, F., Aide, T.M., Zambrano, A.M.A., Balvanera, P., Becknell, J.M., Boukili, V., Brancalion, P.H., Broadbent, E.N., Chazdon, R.L. and Craven, D., 2016. Biomass resilience of Neotropical secondary forests. *Nature*, 530(7589), p.211.
 35. Werner, J.J., Zhou, D., Caporaso, J.G., Knight, R. and Angenent, L.T., 2012. Comparison of Illumina paired-end and single-direction sequencing for microbial 16S rRNA gene amplicon surveys. *The ISME journal*, 6(7), p.1273.

ADDED!

Methods and Materials

Sites and Location. The six sites used in our study are 0.1 ha transects, part of the Agua Salud secondary rainforest plots in the Panama Canal Area (9°N, 79°W) of Central Panama under the supervision of Smithsonian Tropical Research Institute (STRI). Rainforests have developed naturally from the existing seedbank following pasture abandonment. The region most likely consisted of primary rainforests prior their use for pastures. All selected transects were in 17 year old forests with standing biomass similar to that of other secondary forests worldwide³⁴. The forests developed on the same basalt parent bedrock, exhibiting weathered P-poor oxisol soil profiles of considerable (equal to or greater than 15m) depth. Mineralogically, the soils are dominated by kaolinite, goethite and quartz but inclusions of primary minerals from the bedrock (or regolith) brought up by landslides and erosion were infrequently recovered (Supplementary Note 1). The six chosen sites represent a naturally occurring gradient in N₂-fixer basal area (BA) with N₂-fixing legumes occupying 6% (Sites A and B), 9% (Site C), 18% (Site D), 23% (Site E) or 27% (Site F) of total tree BA.

Tree Selection. We selected commonly occurring tree species with the group of N₂-fixers including *Inga cocleensis*, *Inga thibaudiana*, *Abarema barbouriana*, *Swartzia simplex* and *Platymiscium dimorphandrum* – all from the Fabaceae (Leguminosae). The group of non-fixers included representatives of major tropical tree families including *Miconia argentea* (Melastomataceae), *Xylopia frutescens* (Annonaceae), *Terminalia amazonia* (Combretaceae), *Lacistema aggregatum* (Lacistemataceae) and the non-fixing legume *Senna dariensis* (Fabaceae). Each tree species was represented in $n \geq 3$ in all microbiome analyses and in $n \geq 4$ for pooled mineral samples (prior pooling). Specimens were selected by their diameter at breast height (DBH), with median DBH=61 mm and mean DBH=75 mm. Trees were arbitrarily classified by their vertical position as understory, mid-canopy or canopy specimens with fixers and non-fixers having comparable canopy specimen proportions (32% and 35% canopy specimens, respectively).

Soil analysis and Mineral weathering. Four grams of crushed dunite rock (92% olivine; Åheim plant, Norway; grain diameter 250-500 μm / 500-1000 μm = 50 / 50%) were heat-sealed in 5x5x3 cm triangular mesh bags (pore diameter 30 μm). Two such mesh bags were buried 10 cm away from opposite sides of the main trunk at 10 cm depth for a total of four bags per tree. A total of 500 olivine rock bags were buried beneath 125 trees. Soil (0-10 cm depth) was collected from the sites of olivine bag deposition using a sterile hand spade. Collected soils were stored at 4°C for a week and analysed using nitric acid digests, ICP-OES and C/N analyser (for total C and N) at STRI Soil Lab. Sample pH was measured in 0.01 M CaCl₂. After 8-months incubation time, bags were recovered and the contents of all four bags per tree pooled into a single sample of reacted olivine per tree. After taking a subsample for DNA extraction (see below), pooled samples were homogenized to fine powder using an agate ball-mill machine. The nutrient concentrations of resulting finely homogenized powder samples were analysed using portable X-ray fluorescence machine (Niton™ XL2 GOLDD XRF analyser) set at mining mode. Weathering rates were calculated as the change in Mg concentrations relative to fresh unweathered material over time.

DNA extraction, amplicon and shotgun metagenome sequencing. Total DNA was extracted from 0.50 g of each reacted olivine sample pool prior to further homogenization and from 0.25 g of rhizospheric topsoil using the MoBio PowerSoil DNA isolation kit, using the manufacturers' protocols. Amplicon libraries were constructed using a two-step PCR approach where the target region was first amplified using locus specific primers with Illumina sequencing primer adapters incorporated into their 5' end (30 cycles), then a second round of PCR was done using forward and reverse primers that contain all

Illumina sequencing primer and flow cell binding sequences as well as a unique 8 bp barcode (6 cycles). In the first round of PCR, all samples were amplified in triplicate using the 515F/806R primer set that amplifies the V4-V5 region of the 16S rRNA for bacteria³⁵. Triplicate reactions were then pooled and unique combinations of barcodes and Illumina adapters added via a second round of PCR. PCR products were purified and normalized using SequalPrep Normalization plates (Life Technologies), pooled into single libraries based on sample type (e.g. soil rhizospheric 16S, soil mineral 16S), concentrated using Agencourt AMPure XP beads, quantified on a Qubit fluorimeter, and quality checked using the High Sensitivity Agilent DNA kit on an Agilent Bioanalyzer. Subsequently, samples were adjusted to appropriate concentrations and sequenced on a total of two runs on an Illumina MiSeq sequencer (2x250 bp PE runs) at the Smithsonian Tropical Research Institute. Shotgun metagenome libraries were constructed using a total of 15 total DNA samples (12 from reacted mineral and 3 from rhizospheric soil) using Nextera XT DNA Library Prep Kit and following protocols within and sequenced at 1 run of HiSeq 4000 sequencer (2x150 bp PE runs) at the Edinburgh Genomics Institute, UK.

Operational taxonomic unit (OTU) picking and shotgun metagenomic annotation. The resulting 16S rRNA amplicon libraries were submitted for taxonomic analysis through the MG-RAST pipeline¹⁷ and blast2go and the SILVA SSU and NCBI microbial 16S databases were used for taxonomy calling. N, S, Fe cycling activities are assigned to OTUs based on detailed literature searches identifying a total of 6 and 79 genera from the SILVA SSU database with function in the N and S cycle, respectively, and 59 species-level OTUs from the NCBI database with function in the Fe cycle. Shotgun metagenomes were also submitted through the MG-RAST pipeline¹⁷ and annotated according to default characteristics (1e⁻⁵, 40% similarity, representative hit). Annotated metagenomes are publicly available on the MG-RAST server under MG-RAST ID numbers found in Supplementary Note 8. Annotations are based on blat hits against MG-RAST Subsystem and KO databases. Excel spreadsheets containing OTU tables and shotgun metagenomic hierarchies are available as on request. Both metagenomic and 16S RNA hits were expressed in relative abundance dividing their number of reads by the total number of reads. Shotgun metagenome entries were also expressed in relative abundance based on the DNA-directed RNA polymerase beta' subunit (*rpoC*) gene (reads/reads assigned to *rpoC*) to account for different number of genomes per sample as each genome contains a single copy *rpoC*. Further information on methods involved in molecular analyses can be found in Appendices 1-3.

Statistical analyses. All statistical analyses were carried out in the R environment or in GraphPad Prism 7. Sequencing libraries were not rarefied with all libraries (regardless of their number of aligned hits) were included in statistical comparisons. Comparisons between tree groups (N₂-fixers vs. non-fixers) were carried out as two-tailed t-tests or Welch's t-test depending on whether the assumption for equal variances was met. In the case of community analyses, most species-level OTUs exhibit non-normal distribution (Shapiro-Wilk normality tests) and as such comparisons between the microbiomes of N₂-fixers and non-fixers were compared using Mann-Whitney t-tests. Pearson correlation tests were utilized for testing for co-variance between olivine weathering rates and metabolic and candidate genes (P<0.05). For comparisons between 3 or more groups of trees, one-way ANOVA and subsequent Post-Hoc Fisher LSD tests were utilized. PERMANOVA analyses were carried out using the Adonis function in R (dissimilarity matrix method Bray-Curtis, replicates=999). Non-dimensional multiscale (NDMS) principal coordinates analysis (PCoA) analyses were carried out in R using the dissimilarity matrix method Bray. Hierarchical clustering of the correlation heat-map matrix (with each cell

CHANGED!

containing a Pearson r value representative of the relationship between pathways X and Y on axes x and y) was performed in R using Manhattan dissimilarity index and complete clustering method.

Figures and Figure Text

Figure 1. N₂-fixers and nodulation are linked to greater mineral weathering in tropical forests correlated to lower soil C:N ratio and pH. **a.** N₂-fixers reveal doubled weathering rates compared to non-fixers (Welch's t-test, $P < 0.05$); **b.** Non-fixers near N₂-fixers (NF-near) show intermediate olivine weathering rates between non-fixers far from fixers (NF-far) and N₂-fixers (ANOVA, $P < 0.05$); **c.** Soil pH is significantly lower beneath N₂-fixers and NF-near in comparison to NF-far (ANOVA, $P < 0.001$) with similar pattern observed in the pH of weathered olivine; **d.** Soil C:N ratio is significantly lower beneath N₂-fixers and NF-near relative to NF-far from fixers (ANOVA, $P < 0.001$); **e.** Nodulation promotes increasing olivine Ni%, signature to weathering of olivine into secondary clay minerals (two-tailed t-test $P < 0.05$); **f.** Nodulation is the main effect on olivine Ni% rather than the taxonomically-defined ability to fix N₂ (ANOVA, $P < 0.05$). Different letters indicate statistically different means using Fisher's LSD post-hoc test. Error bars reveal SEM.

Figure 2. Metagenomics of the microbial community associated with weathered minerals in tropical forests link increased respiration, nitrogen and carbohydrate metabolic potential of the microbial community to enhanced weathering. **a.** Correlation heatmap matrix of gene abundance allocated to MG-RAST Subsystem-Level 2 pathways for the 12 sequenced metagenomes of reacted olivine at different weathering rates. The heatmap is constructed using R with Manhattan dissimilarity index, complete clustering method and Pearson test correlations (salmon colour indicates positive correlation, blue – negative and white – lack of strong correlation); **b.** Pearson r plot of *rpoC*-normalised MG-RAST Subsystem Level 3 pathways correlating with weathering rates and anaerobic characterisation; The asterisk shows $P < 0.05$; **c.** High level pathways encompassing Respiration, Nitrogen Metabolism and the Krebs cycle all reveal patterns of increase following the order NF-far < NF-near < N₂-fixers; **d.** Metagenomic reconstruction of the microbial N cycle in the mineral showing positive correlations (red asterisks, Pearson test, $* < 0.05$, $^{\wedge} < 0.10$) with weathering rates; glutamine synthetase gene orthologue marked with blue asterisk negatively correlates with weathering as it is most required at low external NH₄⁺ concentrations; **e.** Enrichment in Krebs cycle genes correlating with weathering and links to N cycling revealing evidence for the combined role of high N and C availability. Error bars indicate SEM.

Figure 3. N₂-fixers impact the biogeochemical potential of soil as recorded by enrichment in N, S and Fe cycling lineages. **a.** N₂-fixing legumes reveal enrichment in dominant ammonia oxidising archaeal lineages generating acidity through their conversion of ammonia to nitrite but not in nitrite oxidisers; this enrichment is also evident in NF-near neighbouring fixers relative to NF far from fixers; **b.** N₂-fixers reveal greater cumulative abundance of S cycling genera in their associated soil microbiomes than NF-far linked to higher abundance of several phylogenetically distinct S cycling genera; we also observe community effects of N₂-fixers on non-fixers with NF-near legumes revealing similarly enriched S cycling genera relative to non-fixers far from fixing legumes; **c.** N₂-fixers and NF-near exhibit higher cumulative abundance of Fe(III) reducing genera. Microbial profiles in a-b. are constructed from taxonomy calling of the SILVA SSU database in the MG-RAST environment (with $n=12$, $n=13$, $n=21$ for NF-far, NF-near and N₂-fixers, respectively). The Fe-reducing profiling is constructed using the NCBI 16S microbial database in the Blast2GO environment as outlined in Appendix 1. Statistical tests include Kruskal-Wallis for community effects together with Dunn's post-hoc test and Mann-Whitney test for differences between N₂-fixers and non-fixers without community effects. Error bars show SEM.

CHANGED!

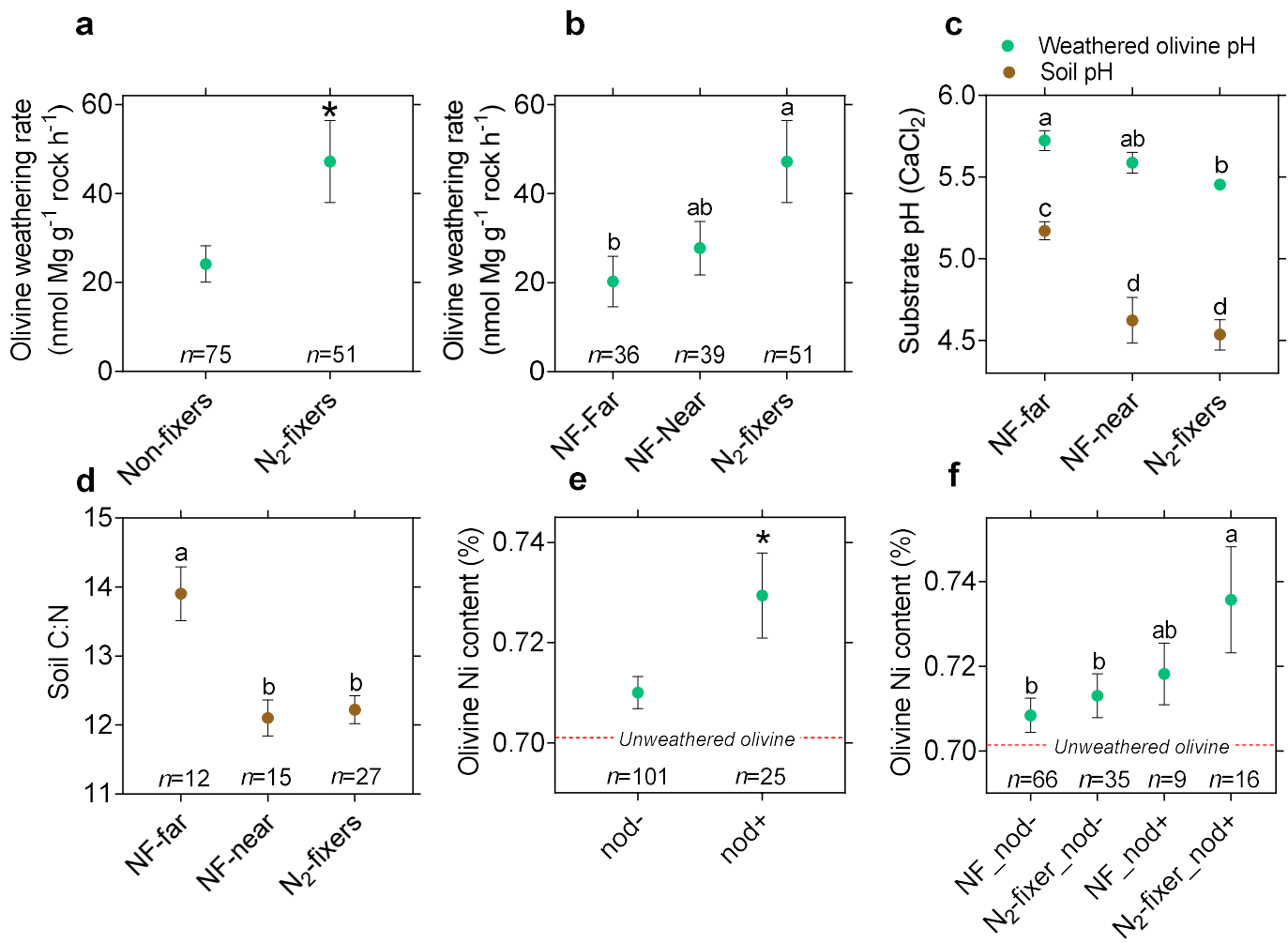
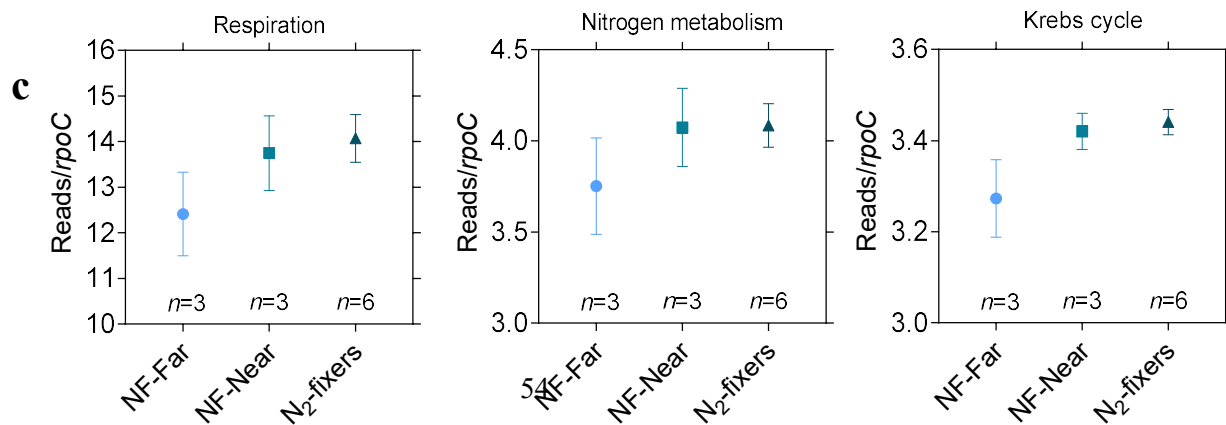
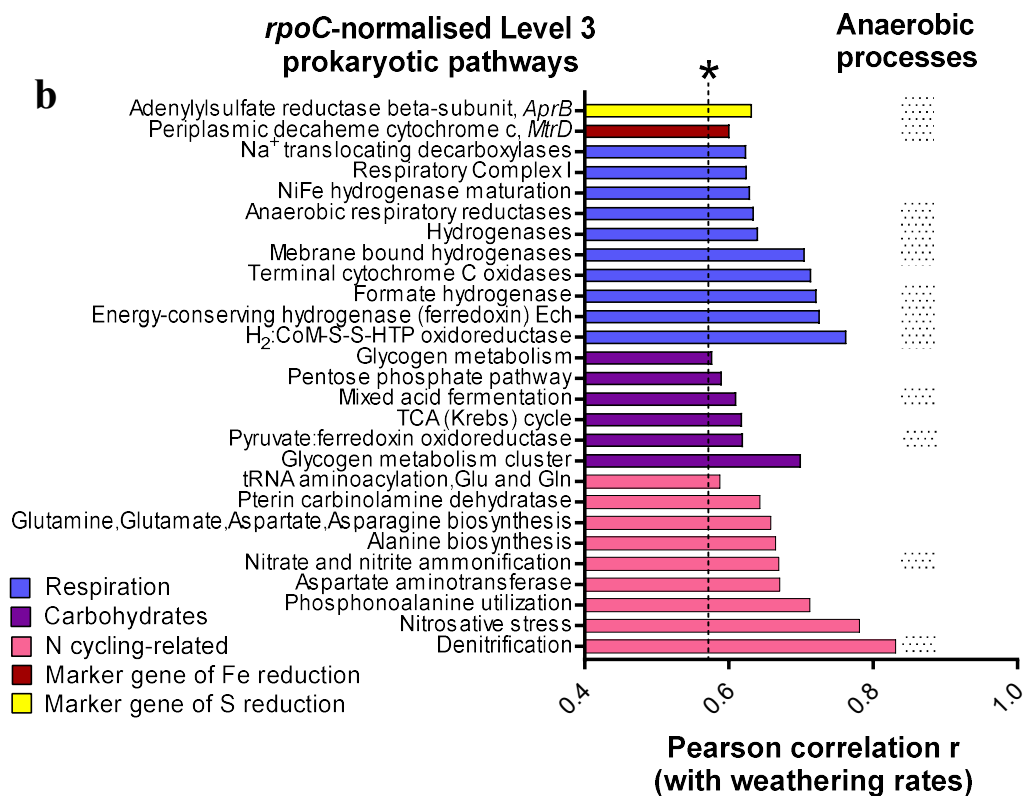
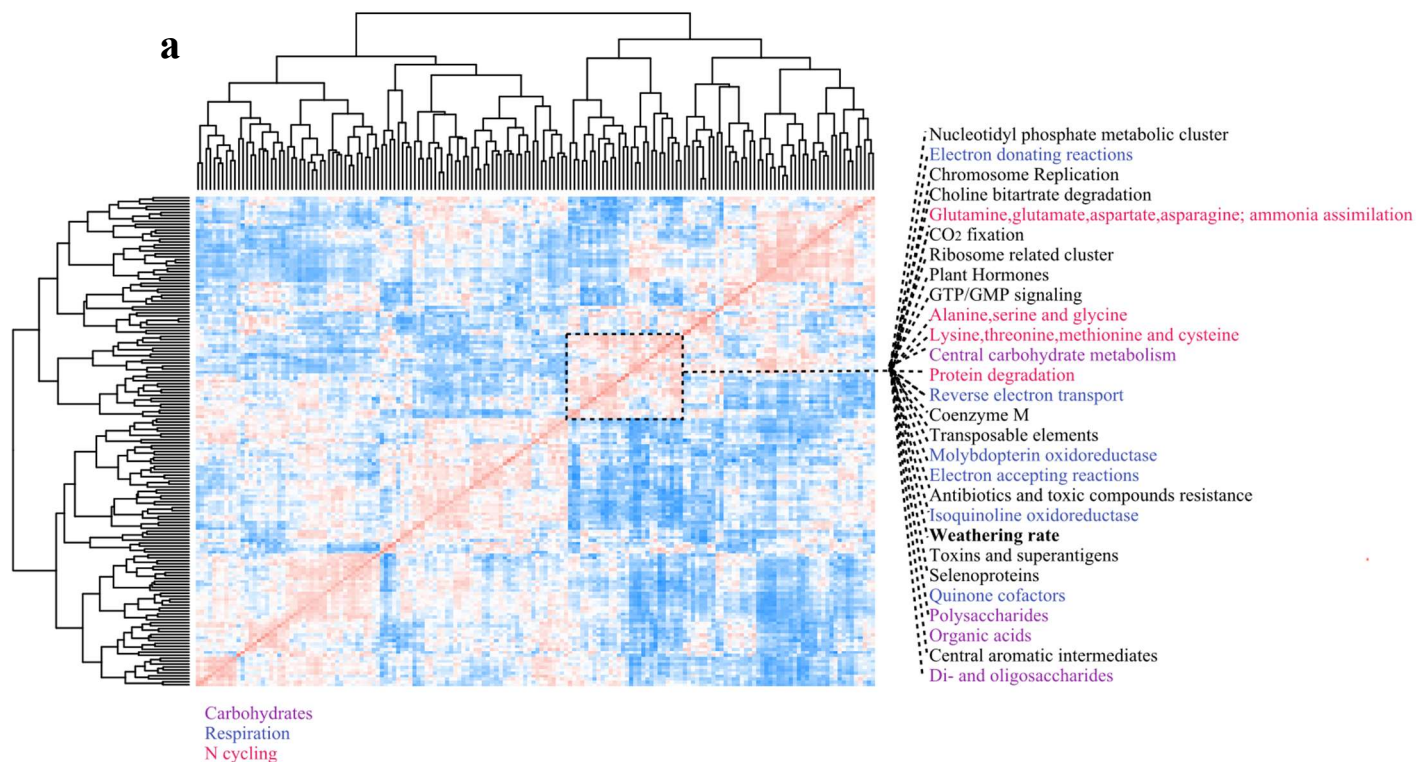


Figure 1. (Epihov *et. al*, 2018)



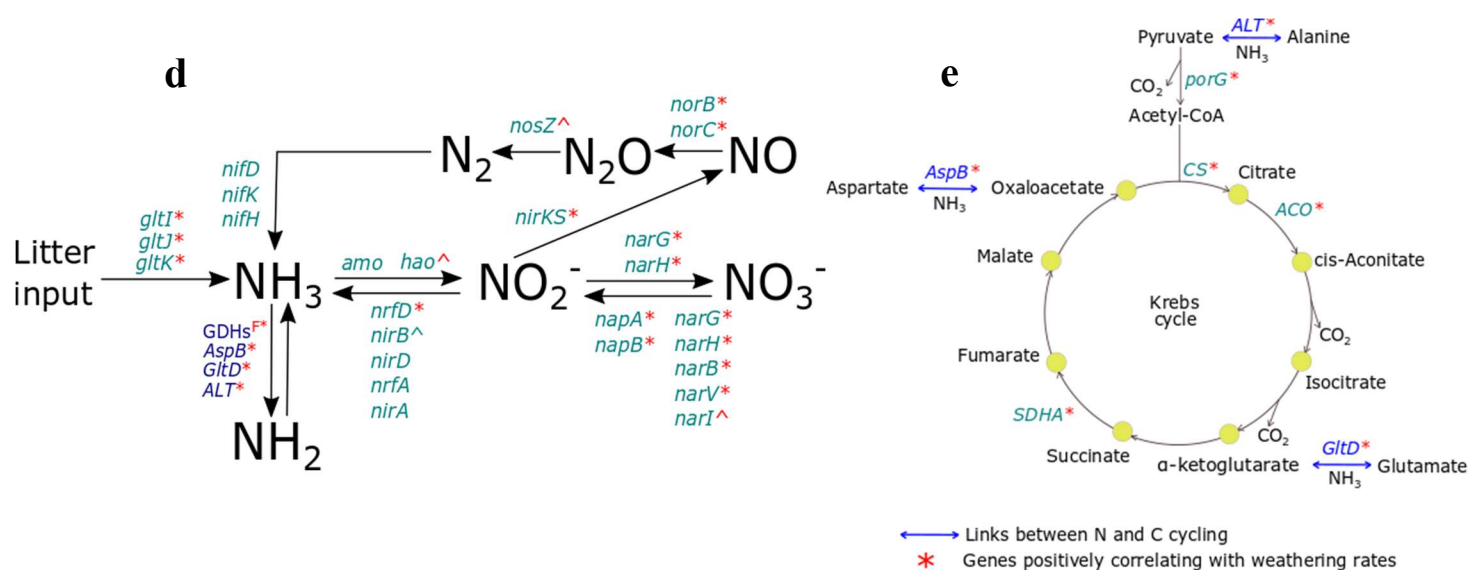


Figure 2. (Epihov *et. al*, 2018)

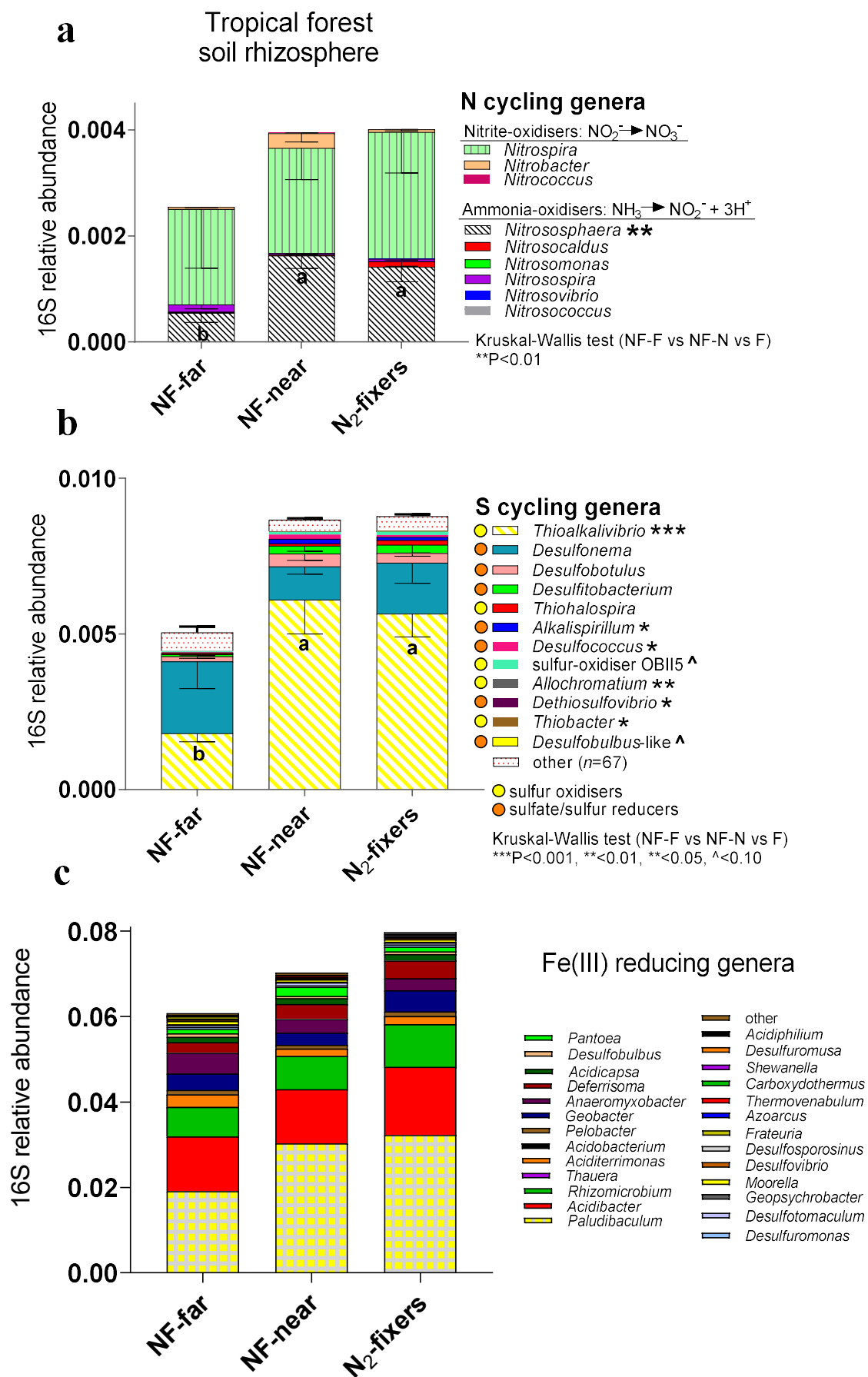


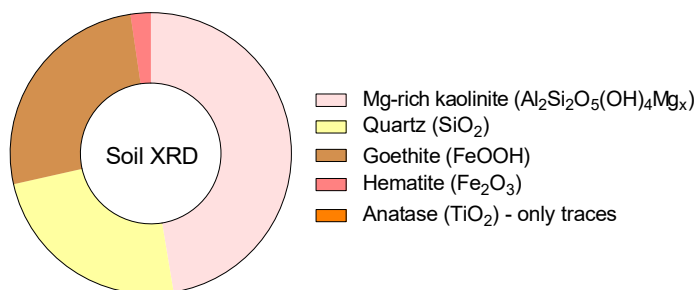
Figure 3. (Epihov *et. al*, 2018)

Supplementary Information

Supplementary Note 1 | Soil chemistry and mineralogy in Panamanian tropical forests

1.1 Soil mineralogy

The average soil mineralogical composition was analysed using X-ray diffraction methods (XRD; X-ray Mineral Services Ltd, Colwyn Bay, UK). The report summarized in Supplementary Figure 1 shows that the main fraction



Supplementary Figure 1. XRD mineralogical profile of Agua Salud rainforest oxisols

is dominated by kaolinite (47.4%) bearing certain structural and chemical resemblance to the Mg-rich chlorite ($\text{Mg}_5\text{Al}(\text{AlSi}_3\text{O}_{10}(\text{OH})_8$), followed by goethite (26.2%), quartz (24.1%), hematite (2.3%), and anatase (traces). The XRD data of our soil samples is comparable to other oxisols

with a characteristic lack of primary minerals[1][2]. However, weathered rock fragments (5-10 cm in length and ~5 cm in diameter) were recovered from several topsoil samples in plot D. Some of those were subsequently crushed and their elemental composition analysed using X-

ray fluorescence (XRF – Supplementary Table 1).

Supplementary Table 1. Mean XRF oxide data (%) comparing soil rock fragments (partially weathered basalt) and fresh basalt rock

Comparison	SiO ₂	MgO	Al ₂ O ₃	FeO	CaO	TiO ₂	SrO	P ₂ O ₅	K ₂ O	CuO	MoO ₃	MnO	ZnO
Soil rock fragments	45.30	10.26 [†]	24.12	15.91	2.70	0.65	0.11	0.01	0.10	0.02	<D	0.20	0.01
Fresh basalt	48.01	10.39	16.75	12.03	6.90	1.79	0.47	0.91	2.08	<D*	<D	0.24	0.02

*<D – below detection limit

[†] - MgO decline in partially weathered soil rock fragments relative to fresh material is probably substantially underestimated due to the formation of secondary minerals between weathered Mg and pedogenic Al (e.g. chlorite of the type clinocllore) covering the partially weathering fragments

n=3 weathered soil fragment from site D, n=3 fresh reference basalt material (from Oregon, USA)

1.2 Soil chemistry and differences between under-crown total soil elements in non-fixers and N₂-fixers

The soils in the Agua Salud secondary rainforests are characterised as weathered oxisols derived from pre-Cenozoic basalt parent material[3][4]. Nitric acid digests and subsequent ICP-OES and C-N analysis (Smithsonian Soil Lab, Ancón, Panama) reveal that total soil C, N and P (Supplementary Table 1) are relatively low compared to similar tropical forest plots in Panama[5], while Cu (particularly in non-fixers – Supplementary

Figure 4) is in the high end of its range[6]. Comparing the levels of Al, Fe and soil cations, we show that N₂-fixing status is a major factor determining lower concentrations of those elements (Supplementary Table 2). The solubilisation and translocation of clays and secondary iron minerals to lower soil horizons fits the observed pattern as during translocation cations also co-migrate with clay explaining the lower levels of almost all elements in N₂-fixers (except K, Na, Mn).

Supplementary Table 2. Differences between total soil elements in under-crown topsoil (0-10 cm) of N₂-fixers and non-fixers and fixer-poor and fixer-rich forests in 17-year old forest sites.

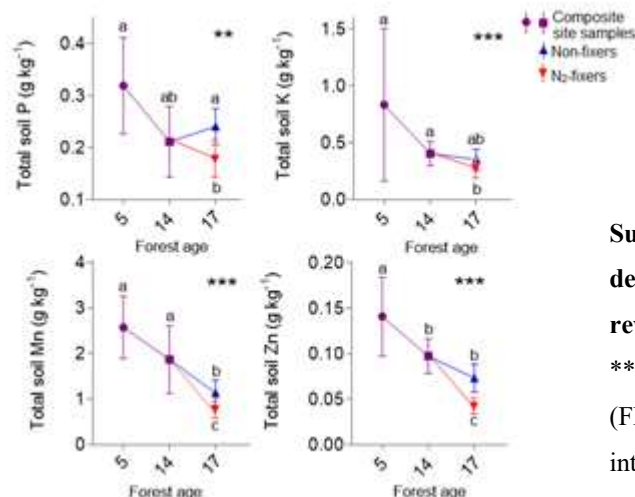
Factor/Element	Al	B	Ca	Cu	Fe	K	Mg	Mn	Na	P	Zn	pH	C	N	C/N	Fe/P
N ₂ -fixing status	**	**	*	***	**	ns	**	ns	ns	**	**	*	ns	ns	ns	ns
Fixer richness	ns	ns	ns	ns	ns	ns	*	ns	*	ns	ns	***	**	*	***	***

Two-way factorial ANOVA type II, ***P<0.001, **P<0.01, *P<0.01, ^{ns}P>0.05. No significant interaction effects were recorded.

Red typeface indicates that the given element is significantly lower in soils of N₂-fixers. Fixer richness differentiates between fixer- or legume-poor forest sites (*n*=3, fixer BA 6-9%) and legume-rich sites (*n*=2, fixer BA 23-27%). Numbers of replicates: *n* fixers=27, *n* non-fixers=27.

1.3. Evidence for decline in total soil P, K, Zn, Mn along the chronosequence

The lower total P in soils of N₂-fixers in 17-year old forests may indicate initial preference and competitive advantage under low P of that functional group during forest establishment in early secondary succession. Alternatively, that pattern may be caused by active decline due to higher solubilisation (release from Al- and Fe-bearing mineral complexes) and increased P uptake and translocation to lower soil horizons. We compared total soil P along the forest chronosequence using site composite samples (each consisting of three pooled 10 cm soil cores from three random locations at sites) for 5- and 14-year old forests. Total soil P declined from 5 to 17 years



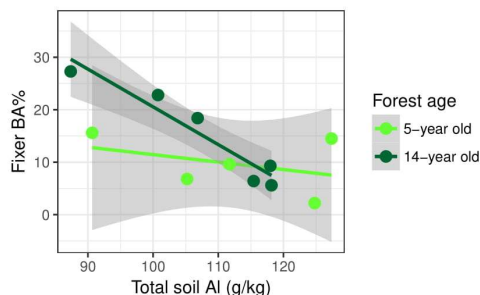
old forests in N₂-fixers but not in non-fixers (Supplementary Figure 2) suggesting that fixers may actively act to reduce P to a greater extent than their non-fixing counterparts. Similarly, K,

Supplementary Figure 2. Total soil P, K, Mn and Zn decrease along the chronosequence with N₂-fixers revealing the lowest levels. ANOVA test, ***P<0.001, **P<0.01, Benjamini-Hochberg False Discovery Rate (FDR) correction. Error bars indicate 95% confidence intervals.

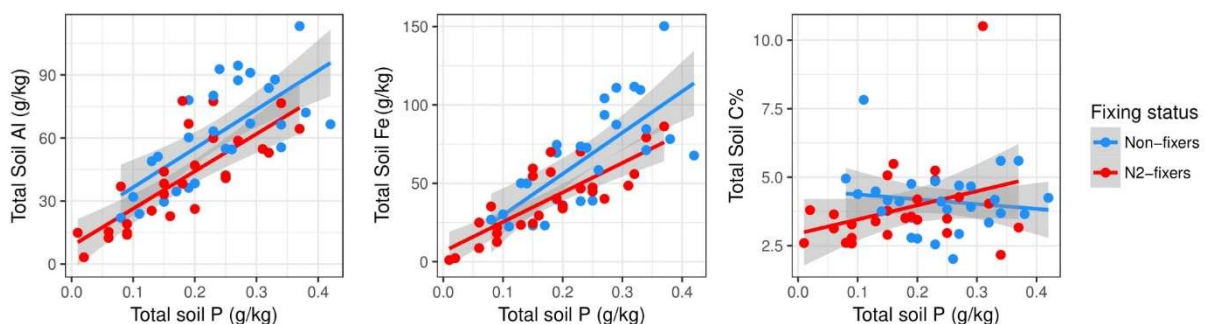
Mn and Zn decreased suggesting that forests development may affect nutrient dynamics. Meta-analysis of the RAINFOR database containing 1388 entries for non-fixers and 148 for N₂-fixers from tropical forests around the globe shows 7.9% higher foliar P in N₂-fixing trees relative to non-fixers (Mann-Whitney test, P<0.01).

1.4. Soil Al, Fe and P and evidence for weathering of Al clays and release of P

Total soil phosphorus strongly correlated with total soil Fe (Pearson correlation test, df=55, P<0.001, r=0.78) and total soil Al (Pearson correlation test, df=55, P<0.001, r=0.77) but not with total soil C% (Pearson correlation test, df=55, P=0.25, r=0.15; Supplementary Figure 3). That suggests that the dominant fraction of P in our soils is not organic but inorganic bound to Al and Fe in either amorphous or crystalline forms (such kaolinite and goethite) – both insoluble and recalcitrant[7]. These results are in line with previous work in Panamanian lowland tropical forests, in which total organic P in topsoil (0-9cm) was only 16% from total soil P, with inorganic phosphorus accounting for the remaining 84%[8]. Interestingly, fixer basal area (BA%) correlate inversely with total soil Al in older sites (14-year old forest sites; Pearson test P<0.01, r=-0.96) but not in young sites (5-year old sites – Supplementary Figure 3). Total soil Al is a poor predictor of available Al³⁺ in the soil solution and thus of Al³⁺ toxicity thus ruling out preference in N₂-fixers to low Al sites due to sensitivity to Al³⁺. Alternatively, soil acidification linked to N₂-fixers during succession (Supplementary Table 2), similar to acid rain effects[9], may cause solubilisation and leaching (translocation) of poorly crystalline Al and

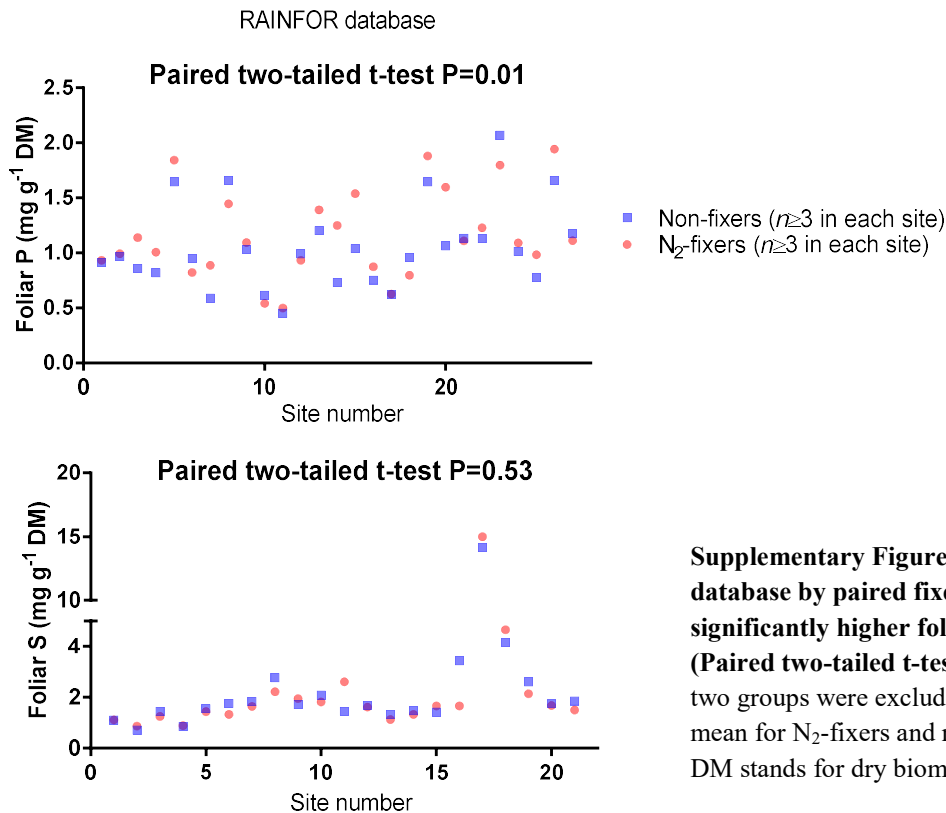


Supplementary Figure 3. Topsoil (0-10cm) total Al inversely correlated with fixer BA% in older forests but not in young forests – all suggesting increased dissolution of Al clays and release of P and correlation between Fe/Al and P but not organic C. Pearson correlation test (see text above for P and r values).



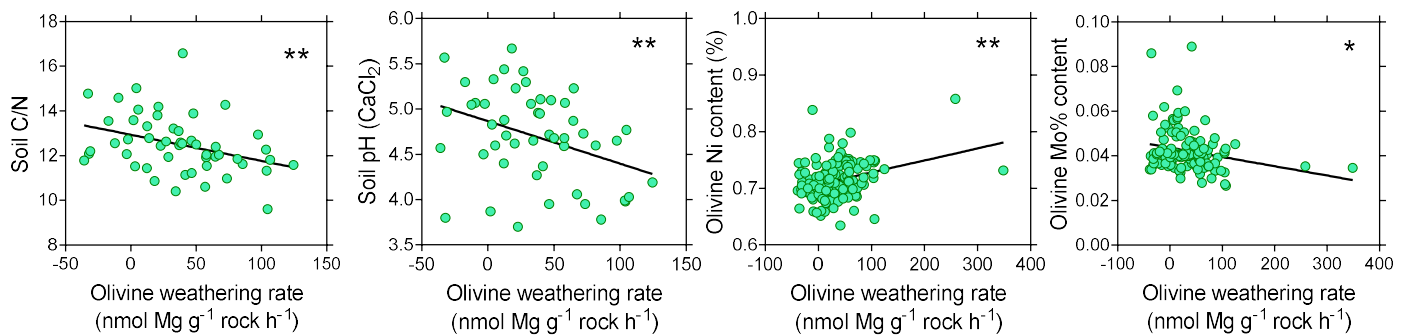
Fe clay mineral phases deeper into the soil thus releasing P. This would help to explain the pattern of decreasing P along the chronosequence in N₂-fixers but not in non-fixers as seen in Supplementary Figure 2. The release of inorganic P may help to explain the greater foliar P levels in N₂-fixers relative to non-fixers observed during our

meta-analysis of the RAINFOR dataset from the TRY database[10] encompassing a total of 1536 entries from tropical forest trees.



Supplementary Figure 4. Meta-analysis of the RAINFOR database by paired fixer/non-fixer site comparisons shows **significantly higher foliar P and no difference in foliar S (Paired two-tailed t-test)**. Sites containing $n < 3$ for each of the two groups were excluded from the analysis. Consequently, the mean for N_2 -fixers and non-fixers for each site is based on $n \geq 3$. DM stands for dry biomass.

Supplementary Note 2 | Spearman correlations between weathering rates and olivine chemistry and soil physicochemical factors



Supplementary Figure 5. Olivine weathering rates correlates with declines in soil pH, soil C: N and enrichment in olivine Ni and depletion in olivine Mo% contents. Spearman test ** $P < 0.01$, * $P < 0.05$.

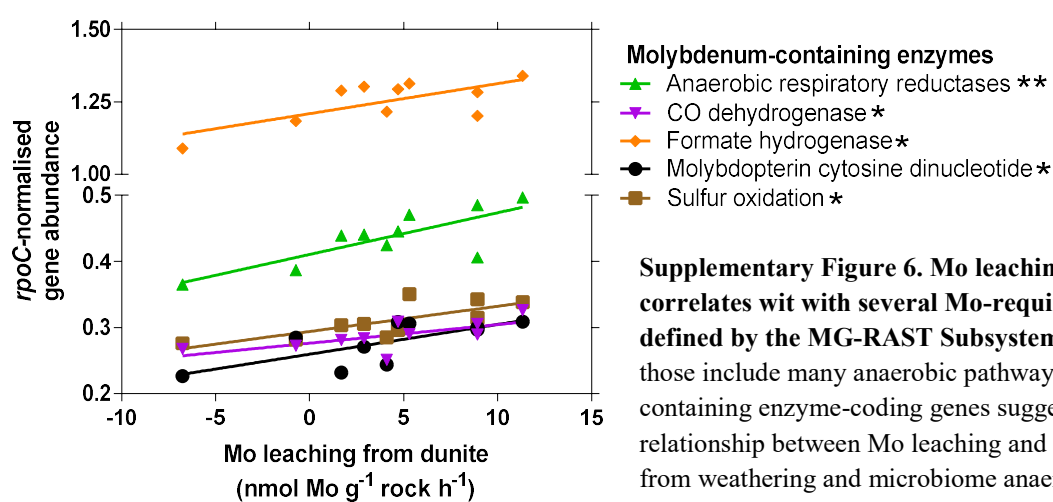
Supplementary Note 3 | Correlations between metagenome gene abundance and weathering rates

Supplementary Table 3. Single gene correlations with weathering rates. The selection includes genes and pathways discussed in the Main Text.

Level 3 pathways enriched in N ₂ -fixers					
P-value	r/rho	Test	Normalisation	Database	Name
0.041596	NA	Mann-Whitney	<i>rpoC</i>	MG-RAST Subsystem	Fructoselysine_(Amadori_product)_utilization_pathway
0.024975	NA	Mann-Whitney	<i>rpoC</i>	MG-RAST Subsystem	Glutamate_dehydrogenases
Gene Orthologues correlating with weathering rates					
0.020815	0.654911	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Citrate synthase (si) (EC 2.3.3.1)
0.045561	0.585327	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Aconitate hydratase (EC 4.2.1.3)
0.029416	0.626066	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Succinate dehydrogenase flavoprotein subunit (EC 1.3.99.1)
0.024998	0.639983	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Pyruvate:ferredoxin oxidoreductase, gamma subunit (EC 1.2.7.1)
0.020035	0.657933	Pearson	Total	MG-RAST KO	E6.4.1.1B, pycB; pyruvate carboxylase subunit B [EC:6.4.1.1]
0.001224	0.815507	Pearson	Total	MG-RAST KO	attA2; mannopine transport system permease protein
0.00122	0.815626	Pearson	Total	MG-RAST KO	attC; mannopine transport system substrate-binding protein
0.033245	0.615179	Pearson	Total	MG-RAST KO	E2.4.1.21, glgA; starch synthase [EC:2.4.1.21]
0.047201	0.581802	Pearson	Total	MG-RAST KO	glgB; 1,4-alpha-glucan branching enzyme [EC:2.4.1.18]
0.009749	0.709546	Pearson	Total	MG-RAST KO	glgC; glucose-1-phosphate adenyltransferase [EC:2.7.7.27]
0.013754	0.6861	Pearson	<i>rpoC</i>	MG-RAST Subsystem	1,4-alpha-glucan (glycogen) branching enzyme, GH-13-type (EC 2.4.1.18)
0.021	0.654207	Pearson	<i>rpoC</i>	MG-RAST Subsystem	glgP; Glycogen phosphorylase (EC 2.4.1.1)
0.029738	0.625113	Pearson	<i>rpoC</i>	MG-RAST Subsystem	talA; Transaldolase (EC 2.2.1.2)
0.020811	0.654928	Pearson	<i>rpoC</i>	MG-RAST Subsystem	kdgK; 2-dehydro-3-deoxygluconate kinase (EC 2.7.1.45)
0.002729	0.780666	Pearson	Total	MG-RAST KO	PGLS; pgl; 6-phosphogluconolactonase [EC:3.1.1.31]
0.046936	0.582367	Pearson	<i>rpoC</i>	MG-RAST Subsystem	xfp; Fructose-6-phosphate phosphoketolase (EC 4.1.2.22)
0.037444	0.604229	Pearson	<i>rpoC</i>	MG-RAST Subsystem	xfp; Xylulose-5-phosphate phosphoketolase (EC 4.1.2.9)
0.046052	0.584263	Pearson	Total	MG-RAST Subsystem	pelW; Exopolysaccharuronate lyase (EC 4.2.2.9)
0.000426	0.852589	Pearson	Total	MG-RAST Subsystem	uxaA; Altronate hydrolase (EC 4.2.1.7)

0.04819	0.57972	Pearson	Total	MG-RAST Subsystem	exuR; Hexuronate utilization operon transcriptional repressor ExuR
0.038451	0.601735	Pearson	<i>rpoC</i>	MG-RAST Subsystem	GltD; Glutamate synthase [NADPH] small chain (EC 1.4.1.13)
0.016861	0.671195	Pearson	<i>rpoC</i>	MG-RAST Subsystem	ALT (alanine transaminase NCBI search); Uncharacterized PLP-dependent aminotransferase YfdZ
0.024831	0.64054	Pearson	<i>rpoC</i>	MG-RAST Subsystem	AspB (closest match in nBLAST); Phosphonoalanine aminotransferase
0.000303	0.862786	Pearson	<i>rpoC</i>	MG-RAST Subsystem	FdoH; Putative formate dehydrogenase iron-sulfur subunit (EC 1.2.1.2)
0.017208	0.669661	Pearson	<i>rpoC</i>	MG-RAST Subsystem	FdhF; tungsten-containing formate dehydrogenase alpha subunit
0.036184	0.607419	Pearson	<i>rpoC</i>	MG-RAST Subsystem	FdhO; Formate dehydrogenase-O, iron-sulfur subunit (EC 1.2.1.2)
0.030197	0.623767	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Anaerobic dehydrogenases, typically selenocysteine-containing
0.029025	0.627235	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Fumarate reductase subunit C
0.014033	0.684664	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Molybdopterin oxidoreductase (EC 1.2.7.-)
0.041476	0.594512	Pearson	Total	MG-RAST Subsystem	Glycerol dehydrogenase (EC 1.1.1.6)
0.024851	0.640473	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Pyruvate formate-lyase (EC 2.3.1.54)
0.031527	0.619947	Pearson	<i>rpoC</i>	MG-RAST Subsystem	quinoprotein alcohol dehydrogenase
0.019298	0.660868	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Formate hydrogenlyase subunit 3
0.012214	0.694424	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Formate hydrogenlyase subunit 4
0.042004	0.593288	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Formate hydrogenlyase subunit 7
0.044197	0.588326	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Formate hydrogenlyase transcriptional activator
0.017293	0.669289	Pearson	<i>rpoC</i>	MG-RAST Subsystem	L-lactate dehydrogenase (EC 1.1.1.27)
0.013824	0.685735	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Phosphate butyryltransferase (EC 2.3.1.19)
0.034844	0.610899	Pearson	<i>rpoC</i>	MG-RAST Subsystem	3-hydroxybutyrate dehydrogenase (EC 1.1.1.30)
0.005391	0.745421	Pearson	<i>rpoC</i>	MG-RAST Subsystem	CoB--CoM heterodisulfide reductase subunit A (EC 1.8.98.1)
0.020159	0.657447	Pearson	<i>rpoC</i>	MG-RAST Subsystem	CoB--CoM heterodisulfide reductase subunit B (EC 1.8.98.1)

0.034531	0.611725	Pearson	<i>rpoC</i>	MG-RAST Subsystem	CoB--CoM heterodisulfide reductase subunit C (EC 1.8.98.1)
0.001814	0.799218	Pearson	<i>rpoC</i>	MG-RAST Subsystem	CoB--CoM heterodisulfide reductase subunit D (EC 1.8.98.1)
0.048978	0.57808	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Coenzyme F(420)H(2) dehydrogenase (methanophenazine) subunit FpoN
0.007409	0.726831	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Energy-conserving hydrogenase (ferredoxin), subunit A
0.017534	0.66824	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Energy-conserving hydrogenase (ferredoxin), subunit B
0.003591	0.76715	Pearson	<i>rpoC</i>	MG-RAST Subsystem	Energy-conserving hydrogenase (ferredoxin), subunit E



Supplementary References

1. Fontes, M.P.F. and Weed, S.B., 1991. Iron oxides in selected Brazilian oxisols: I. Mineralogy. *Soil Science Society of America Journal*, 55(4), pp.1143-1149.
2. Van Wambeke, A., Eswaran, H., Herbillon, A.J. and Comerma, J., 1983. Oxisols. In *Developments in Soil Science* (Vol. 11, pp. 325-354). Elsevier.
3. Batterman, S.A., Hedin, L.O., Van Breugel, M., Ransijn, J., Craven, D.J. and Hall, J.S., 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature*, 502(7470), p.224.
4. Neumann-Cosel, L., Zimmermann, B., Hall, J.S., van Breugel, M. and Elsenbeer, H., 2011. Soil carbon dynamics under young tropical secondary forests on former pastures—A case study from Panama. *Forest ecology and management*, 261(10), pp.1625-1633.
5. Turner, B.L. and Engelbrecht, B.M., 2011. Soil organic phosphorus in lowland tropical rain forests. *Biogeochemistry*, 103(1-3), pp.297-315.
6. Wild, A., 1988. Russet's Soil Conditions and Plant Growth, ELBS.
7. Chacon, N., Silver, W.L., Dubinsky, E.A. and Cusack, D.F., 2006. Iron reduction and soil phosphorus solubilization in humid tropical forests soils: the roles of labile carbon pools and an electron shuttle compound. *Biogeochemistry*, 78(1), pp.67-84.
8. Dieter, D., Elsenbeer, H. and Turner, B.L., 2010. Phosphorus fractionation in lowland tropical

rainforest soils in central Panama. *Catena*, 82(2), pp.118-125.

9. Mulder, J., Van Grinsven, J.J.M. and Van Breemen, N., 1987. Impacts of Acid Atmospheric Deposition on Woodland Soils in the Netherlands: III. Aluminum Chemistry 1. *Soil Science Society of America Journal*, 51(6), pp.1640-1646.

10. Kattge, J., Diaz, S., Lavorel, S., Prentice, I.C., Leadley, P., Bönsch, G., Garnier, E., Westoby, M., Reich, P.B., Wright, I.J. and Cornelissen, J.H.C., 2011. TRY—a global database of plant traits. *Global change biology*, 17(9), pp.2905-2935.

Chapter 2

Multiple symbiont recruitments and nitrogen feedbacks govern enhanced mineral access in monodominant *Acacia* of Australian tropical forests

Dimitar Z. Epihov^{1*}, Alexander W. Cheesman², Lucas A. Cernusak², Miriam Goosem², Susan G. W. Laurance^{2,3}, Michael I. Bird², Jonathan R. Leake¹ and David J. Beerling¹

¹*Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK*

²*College of Science and Engineering and Centre for Tropical Environmental and Sustainability Science, James Cook University, Cairns, Australia*

³*Wet Tropics Management Authority, PO Box 2050, Cairns 4870, Australia*

*for correspondence

Abstract

N₂-fixing legumes are abundant and functionally important for alleviating N limitation during the re-growth of secondary forests^{1,2,3}. *Acacia celsa* is a pioneering N₂-fixing legume in the Australian Wet Tropics area often forming nearly monodominant forest stands⁴. However, what allows *Acacia* to form such communities relative to other non-fixing pioneers as well as the exact contribution of its symbioses are unclear. Here, we utilize a forest chronosequence (12-48 years old *Acacia* stands and mature forests) and show that the transient peak in nodulation and symbiotic N₂-fixation at 20 years of forest re-growth coincides with mirroring peaks in the weathering rates of basalt and dunite silicate rocks with *Acacia* exhibiting significantly higher total weathering. Next-Generation sequencing of 34 full shotgun soil-mineral metagenomes showed significant increase in N cycling pathways upon high nodulation regimes. Consequently, correlational clustering analyses placed weathering enhancement together with multiple N cycling pathways linking fixer-driven changes in the N cycle to parallel increase in acidifying and redox dissolutive processes of the microbial carbohydrate, respiratory and anaerobic metabolism. *Acacia* rhizospheres also had the highest rate of P and K leaching from basalt throughout the chronosequence as well as significantly greater abundance of ectomycorrhizal fungi in roots and basalt communities. Our findings corroborate that patterns in mineral weathering in secondary succession are

temporally-specific and that *Acacia* trees affect weathering both directly (through mycorrhizal recruitment and root activity) and indirectly (through N-driven cascading effects on the free-living belowground microbiome) creating an improved supply of essential nutrients likely contributing to their monodominance in Australian Tropics and implicating the use of their fast growth and their characteristic microbial ecology in future reforestation efforts and sustainable forestry management.

Introduction

Tropical forests around the world are currently experiencing great reductions in area with deforestation activities advancing at rates of 9-12 million hectares per annum or 0.5% total loss of forest cover across the tropics^{5,6}. Rates of naturally occurring forest re-growth during secondary succession can compensate for and recover 20-85% of the loss rate in forest area⁷. Reforestation also acts to increase the average size of forest fragments by 40% and reduce fragmentation (total number of forest fragments) by 15% globally⁶ showing the potential of fragments to eventually coalesce and form clusters of continuous forest cover.

The role of N₂-fixing legume trees during this successional process is well established^{1,2}. Young recovering forests are frequently limited by N creating conditions⁸ favouring the recruitment and high fixation rates of N₂-fixers¹. Also, N₂-fixers in such young N-limited re-growth forests are linked to higher mineral weathering accommodating their higher growth rates than non-fixers. Many such legume-rich communities represent a varied mixture of different N₂-fixing legume species^{1,9} with early pioneering legumes typically receding in abundance in older, more mature forests once P limitation takes over N limitation⁸. In Australia, however, many such young secondary forests are dominated by a single legume species – *Acacia celsa*. *Acacia*-dominated forests were suspected to restrict forest advancement into mature forests but recent investigations show that they do in fact actively recruit late successional species provided there is a nearby fragment of mature forest acting as a seedbank⁴. However, the longevity of *Acacia* trees in such tropical forests assures persistence well into older forests with predicted rates indicating that succession into mature forest may take more than 200 years⁴. We hypothesized that this unusual for pioneering N₂-fixing legumes persistence may be invoked by *Acacia celsa* ability to form multipartite symbioses including partnerships with rhizobial bacteria, arbuscular mycorrhizal (AM) and

ectomycorrhizal (EM) fungi¹⁰ acting to alleviate N and P limitation and allowing access to nutrients derived from mineral weathering boosting *Acacia* growth and fitness relative to other non-fixing AM pioneering trees.

Here, we test this hypothesis in a secondary forest chronosequence in the Atherton Tableland, Queensland, Australia by investigating the growth dynamics and weathering rates and preferential leaching of limiting nutrients by depositing fresh rock material in the rhizospheres of *Acacia* and non-fixing trees along the chronosequence. Subsequently, we investigate the drivers of changes in weathering dynamics by coordinated analysis of (1) soil chemistry ($n=54$), (2) root and rock symbiotic status (through *de novo* assembly of root transcriptomes, $n=11$ and rock eukaryotic large ribosomal subunit metagenomic pulling and taxonomy assignment, $n=34$), (3) the functional profile of free-living microbiota in rock (by shotgun metagenomics and metabolic pathway assignment, $n=34$).

Results and Discussion

A. celsa exhibits faster growth than non-fixers and persistence into older secondary forests

We used inventory data from 26 secondary forest sites encompassing species-level identified counts of over 2,700 tree stems and their diameter at breast height (DBH) forming a chronosequence (9-69 years post-abandonment) to establish the growth dynamics of the 5 target tree species in our study. *Acacia celsa*, the major N₂-fixer in these forests, grew at rates much faster than those observed for any of the target non-fixing species and faster than that of all non-fixers combined (**Figure 1**) allowing it reach an average diameter of 45.3 cm compared to 12.1 cm average for non-fixers in 48-year old forests. It is this high growth rate that supports near monodominance with some 15-30-year old and 30-48-year old forests displaying up to 57% to 64% of their basal area (BA) occupied by *A. celsa*, respectively (**Table 1**) – a range consistent with other cases of monodominant tropical forests¹¹. However, the lack of trees of small DBH in older forests suggesting apparently absent seedling recruitment (**Figure 1a**), also reported in previous work⁴, suggests that unlike mature monodominant forests¹¹, the monodominance of *A. celsa* is a transient successional event¹² carrying into older secondary but not into mature forests.

Age and tree symbiotic characteristics affect weathering patterns

We deposited fresh samples of crushed (250-500 μm grain diameter) dunite and basalt silicate rocks enclosed in polyethylene mesh bags in the rhizosphere (10 cm depth) of 59 trees along the secondary forest chronosequence and under 18 trees in mature forests. After 10 months including a full wet season, bags were recovered and reacted rocks analysed for weathering rates relative to fresh material using X-ray fluorescence. The greatest declines in dunite Mg content and basalt product of weathering index (PWI)¹³ were found in younger (12.5 and 20-year old) forests indicating intense total weathering, with rock weathering significantly decreasing in older successional (48-year old) and mature (120-year old) forests for both types of silicate rocks (**Figure 2a,b**; ANOVA, $P < 0.001$ for both rock types).

The lack of age replication makes it difficult to exclude the role of possible site effects resulting from differing soil conditions and vegetation composition. We carried out soil analyses and non-metric multidimensional analyses using all measured total elements showing that soil conditions are an unlikely driver of the observed differences (**Supplementary Figure 1**). Using vegetation composition data from sites of different ages (12, 5, 20, 48-years old and mature forests, $n=3, 4, 4, 8$, respectively) and Chao clustering, we find that our target sites conform well with successional vegetation differences between forests of different age (**Supplementary Figure 1**), suggesting that differences in weathering likely result from genuine age-related successional changes.

P leaching from basalt did not correlate with overall basalt weathering rates (measured by PWI) revealing incongruent (independent) leaching rates that were highest in 48-year old forests (**Figure 2c**). Minerals beneath *A. celsa* revealed significantly greater P decline than non-fixers suggesting higher P leaching rates (Mann-Whitney test, $P < 0.05$; **Figure 2d**). Similarly, reacted basalt from the rhizosphere of this legume tree species also consistently exhibited significantly lower K levels (Mann-Whitney, $P < 0.01$; **Figure 2d**). The highest weathering rates recorded for both dunite and basalt were those of mineral material beneath N_2 -fixers in the 20-year old forest site (**Figure 2a,b**). Interestingly, they also exhibited the highest levels of nodulation (**Figure 3a**) offering an intriguing link between symbiotic N_2 -fixation, N cycling and weathering dynamics.

Symbiotic N₂-fixation transitions the soil-mineral microbiome to enhanced weathering through cascading N input effects

To further investigate the role of symbiotic N₂-fixation on soil function and weathering, we extracted and sequenced 34 metagenomes from the soil-buried reacted minerals (subsequently referred to as soil-mineral metagenomes). Genes assigned to the N₂-fixation pathways of free-living microorganisms peaked in soil-mineral metagenomes from the 20-year old forests where nodulation was highest (**Figure 3a,b; Supplementary Figure 2**). Using root transcriptomics, we show that nodulated roots contained 4-fold greater abundance of *Bradyrhizobium* transcripts than non-nodulated roots (ANOVA, $P < 0.01$; **Figure 4a**) confirming *Bradyrhizobium* as the main nodulating agent in *Acacia celsa* in agreement with previous nodulating reports for the closely related *Acacia mangium*¹⁴. Over 50% of all Fe-Mo nitrogenase beta subunit (*nifK*) gene copies in basalt were assigned to *Bradyrhizobium* with basalt beneath N₂-fixers containing significantly greater bradyrhizobial *nifK* gene abundance than non-fixers (**Figure 4b**, two-tailed t-test, $P < 0.05$). These findings corroborate that the peak in N₂-fixation pathways in basalt and dunite of 20-year old forests is contributable to high nodulation rates and abundance of bradyrhizobia in the soil of these nodulating trees. Previous studies have found no relationship between *Acacia* tree diameter and abundance of unique reads assigned to the Bradyrhizobiaceae³². Our findings support this relationship by asserting that peak in bradyrhizobial *nifK* reads and N₂-fixation genes occurs not in the largest 48-year old trees but in 20-year old trees suggesting that not tree diameter but N availability and tree-soil feedback relationships may control nodulation and niche construction instead. In the absence of compatible hosts, studies have revealed different species as dominants of free-living N₂-fixing communities including *Burkholderia*, *Geobacter*, *Azotobacter*, *Rhodopseudomonas*, *Methylocella* and *Bradyrhizobium*³³. However, *Bradyrhizobium* dominance of free-living communities in the absence of host is associated with neutral soils³³ and not the highly acidic oxisols present at our sites suggesting that symbiotic attraction is an important driver of bradyrhizobial abundance at our sites.

In addition to this symbiosis-driven effects on the microbial community, high N inputs resulting from legume litterfall rich in fixed N can also impact the microbiome. Nodulation co-occurred with mirroring increases in the abundance of ammonia monooxygenase gene in the soil-mineral metagenomes (**Figure 3a**). Ammonia monooxygenase is a key member of the microbial N cycle catalysing the oxidation of ammonia to nitrite, a process that generates

acidity and responds to high N loads suggesting up-regulation of N cycling activity in this highly nodulating forest.

To further dissect the apparent relationship between N inputs, N cycling and weathering dynamics, we constructed detailed Spearman correlation heat-map matrices for dunite and basalt soil-mineral metagenomes including all of the over 1,100 lower level metabolic pathways and weathering rates. Using complete Manhattan clustering, we show that weathering rates fell into a well-defined cluster of positive correlations (**Figure 3b**). The analysis identified two clusters 1 and 2 (*sensu stricto*) with similarly large positive Spearman rho correlation values with weathering (**Figure 3c**). Cluster 3 pathways also exhibited positive albeit significantly lower rho values (*sensu lato* cluster). In contrast, cluster 4 pathways were characterised by negative correlations with weathering. To establish the shared pathways linked to high weathering rates in both dunite and basalt metagenomes, we selected pathways found in clusters 1 and 2 from both rocks (a pathway was still considered if it was present in cluster 3 provided it was present in cluster 1 or 2 from the other rock).

Our cluster analysis highlighted a total of 17 pathways involved in N transformations and cycling (**Figure 3d**) that positively linked to weathering in both dunite and basalt metagenomes including ammonia monooxygenase (NH_3 to NO_2^-), denitrification ($\text{NO}_2^-/\text{NO}_3^-$ to N_2), dissimilatory nitrite reductase ($\text{NO}_3^-/\text{NO}_2^-$ to NH_3) and glutamate dehydrogenases ($\text{NH}_3 + \alpha\text{-ketoglutarate}$ to glutamate). Metagenomically, high N transforming microbial activity is linked to high N inputs¹⁵ often accompanied by nitrate leaching. As a consequence of nitrate leaching, protons generated by nitrification cannot be compensated by OH^- extrusion coupled to microbial and plant nitrate uptake, eventually causing the build-up of acidity¹⁶ which can stimulate weathering rates.

Also in the positively correlating clusters are pathways involved in carbohydrate metabolism. Of particular interest are the adapted to high C availability Entner-Doudoroff glycolytic pathway¹⁷ and those pathways converting the labile C source glucose to the weathering potent gluconic and 2-ketogluconic acids^{18,19} including genes encoding the enzymes pyrroloquinoline quinone (PQQ)-dependent glucose dehydrogenase and gluconate 2-dehydrogenase (**Figure 3d**). Another pathway sensitive to high C availability positively linked to weathering is that of glycogen synthesis and metabolism – a reserve polysaccharide in bacteria used to store C at high external supply²⁰. Utilisation of maltose, a disaccharide

resulting from decomposition of starch, was also positively linked to weathering and other pathways in clusters 1 and 2 (**Figure 3d**). Together, these results paint a picture of high labile C loading resulting from either enhanced forest productivity in such forests of high fixed N supply or from rapid decomposition of the N-rich legume litter ultimately augmenting microbial activity and respiration.

Microbial respiration generates by-product CO₂, increases its concentration at the soil-mineral interface and stimulates weathering through formation of carbonic acid (H₂CO₃). In support, the gene encoding carbonic anhydrase (catalysing the reversible conversion of CO₂ to H₂CO₃) in basalt metagenomes directly correlates with pathways encoding respiratory dehydrogenases, PQQ-dependent glucose dehydrogenase and N conversions (**Supplementary Table 1**) highlighting the relationship between microbial N, C cycling and respiratory CO₂ evolution. In further support of this relationship, the major respiratory pathway components (respiratory complex I, CO dehydrogenase, H₂:CoM-S-S-HTP oxidoreductase, cytochrome B6-F complex, terminal cytochrome C oxidases) are also contained within the positively correlating clusters of basalt and dunite metagenomes identified previously (**Figure 3d**).

High respiration rates would ultimately drive a decline in available O₂ predicted to create multiple anaerobic microsites. In agreement, pathways involved in anaerobic microbial activities including fermentative processes (lactate fermentation, pyruvate:ferredoxin oxidoreductase), S reduction (sulfate reduction-associated complexes), Fe respiration (Fe(III) respiration – *Shewanella* type), methanogenesis (coenzyme F420-H₂ dehydrogenase methanophenazine, particulate methane monooxygenase pMMO, methanopterin biosynthesis) are also found within the *sensu stricto* positively correlating weathering clusters of both dunite and basalt metagenomes (**Figure 3d**). The anaerobic food chain found in such microsites can be conceived to generate large quantities of fermentative organic acids including acetic, lactic, butyric, formic acids – all sources of acidifying and complexing activities towards enhanced weathering. Furthermore, S reduction and Fe respiration can both drive the redox dissolution of basaltic and dunitic Fe(III) to Fe(II)S or Fe(II), respectively, thus further corroding their consisting minerals^{21,22,23}.

The complex relationship between high N inputs of low C/N ratio driving faster decomposition, enhanced microbial respiration and carbohydrate cycling which then reduces

available O₂ favouring anaerobic metabolism can be described as cascading N input effects on the belowground microbiome function all separately contributing to enhanced weathering (**Figure 3d**). Interestingly, we have first hypothesized the existence of N-driven weathering feedback effects in regards to the impact that the evolution of legume-rich tropical forests had on weathering during the Cenozoic²⁴. Then, we provided the first evidence that the hypothesized cascading N effects were observed in N₂-fixing legume trees of Neotropical forests in Panama beneath which soil organic matter had lower C/N ratio, consistent with elevated N inputs, and weathering rates of dunite (olivine) were significantly higher than beneath non-fixing trees. Our combined findings from Panamanian forests reported before from Australian tropical forests reported here, suggest that the ability of symbiotic N₂-fixing legume trees to impact belowground microbial communities and increase rock weathering through N inputs may be widespread.

Multiple symbiotic recruitments in *A. celsa* provide improved access to basalt nutrients

We assembled *de novo* root transcriptomes from both *Acacia celsa* (*n*=7) and *Alphitonia petriei* (*n*=3) using roots near the weathered rock samples (extracted within the same soil core as the collected weathered rock samples). In addition to the symbiotic recruitment of N₂-fixing nodulating *Bradyrhizobium* spp. (**Figure 4a,b**; ANOVA, *P*<0.01 and two-tailed t-test *P*<0.05), roots and basalt communities of *A. celsa* also demonstrated higher recruitment of ectomycorrhizal fungi (**Figure 4c,d**; Mann-Whitney test, *P*<0.05). While no significant difference was found between fixers and non-fixers in their relative abundance of arbuscular mycorrhizal fungi, basalt metagenomes from *Acacia* contained a significantly greater cumulative abundance (sum) of mycorrhizal fungi (Mann-Whitney test, *P*>0.05; **Figure 4d**). Importantly, all 11 EM fungal genera found in the soil-mineral metagenomes of weathered basalt were also present in root transcriptomes supporting the existence of active physical mycorrhizal connections (links) between roots and rocks (**Figure 4g**).

A significantly greater abundance of ectomycorrhizal fungal LSUs was found in metagenomes of the P and K-containing weathered basalt (*n*=17) than in the P and K-lacking dunite samples (*n*=17, Wilcoxon's paired ranked test, *P*<0.05; **Figure 4e**) providing a tantalizing link between basalt P and K leaching and ectomycorrhization, particularly in the light of the greater P and K leaching observed from basalt beneath *Acacia* trees (**Figure 2d**). Furthermore, we report that samples that had exhibited high P leaching rates were

significantly enriched in the ectomycorrhizal *Pisolithus* and *Hebeloma* (Mann-Whitney test $P < 0.01$ and $P < 0.05$, respectively; **Figure 4h**), implicating these fungal genera in P leaching and transfer to plant roots.

In support, transcriptome analysis of roots extracted from rooting zones of high basalt P leaching exhibited an up-regulated expression of P-dependent metabolic pathways (**Table 2**). Those included carbohydrate phosphorylation, protein phosphorylation, phospholipid translocation, phosphorelay signal transduction system, adenosine triphosphate (ATP) synthesis coupled electron transport (**Table 2**), all signifying a root response to increased bioavailability of phosphorus. Furthermore, we found that high P leaching also correlated with transcript abundance of root defence responses and responses to biotic stimuli, likely outlining the interactions with P-transferring mycorrhizal fungi infecting the roots. Interestingly, proton transmembrane transport, involved in rhizospheric acidification and improved P nutrition, also positively correlated with P leaching (**Table 2**) suggesting an active role of roots in the basalt weathering processes. Similarly, symbiotic N_2 -fixation triggered a characteristic transcriptional response in nodulated roots consistent with improved exchange of N metabolites, protein synthesis and limiting O_2 concentrations (**Supplementary Table 2**).

We find evidence for temporal trade-offs in the symbiotic recruitment patterns using our soil-mineral metagenome libraries. The basalt metagenomes peaked in bradyrhizobial *nifK* sequences in the profusely nodulated 20-year old forest which co-occurred with the lowest counts of mycorrhizal fungi (**Figure 4h**). In contrast, soil-basalt metagenomes from the least nodulating 48-year old forest exhibited the highest peak in ectomycorrhizal LSUs accompanied by a significant drop in bradyrhizobial *nifK*s (**Figure 4h**). Although the lack of replication at the forest age level makes it difficult to definitively differentiate between site and age effects, the observation of timing and trade-offs between nodulation and ectomycorrhization still holds. Previous work with tropical N_2 -fixers shows that symbiotic nodulation is under the control of external N supply, increasing during N-limiting conditions in young secondary forests. N-limitation observed in the initial stages of secondary forest regrowth is often then replaced by P limitation in later successional stages providing a possible explanation for our observed patterns in symbiont abundances. Particularly, tree C photosynthates are only allocated to symbionts alleviating the current nutrient limitation – be it N (through nodulation and diazotrophy) or P (through P-mining ectomycorrhizae). In

further support, soil total P was the at its lowest concentration in the 48-year old forests (Supplementary Figure 1).

Conclusions

We report that *Acacia celsa*, a dominant legume native to secondary tropical forests in Australia, similar to its legume relatives of Panamanian Neotropical forests, regulates its nodulation in likely response to N limitation during secondary forest succession. High nodulation rates co-occur with higher basalt and dunite silicate weathering rates in *Acacia celsa* trees relative to non-fixing trees. We provide soil-mineral metagenomic data revealing that the high inputs of fixed N trigger a suite of cascading N effects resulting in enhanced N cycling, microbial respiration, carbohydrate metabolism and anaerobic activity, all positively clustering with weathering rates and stimulating various weathering mechanisms (redox dissolution, acidification, organic acid generation). Throughout the chronosequence, basalt material from beneath N₂-fixing *Acacia* trees had significantly greater P and K leaching suggesting improved access to these key macronutrients. We report data revealing that in addition to N feedback effects, this ability of legumes may be due to their higher abundance of ectomycorrhizal symbioses in roots and basalt material.

The faster growth rates and monodominance of *Acacia celsa* in secondary tropical forests may be governed by its ability to timely recruit multiple root symbionts with corresponding changes in root expression patterns and microbial community function allowing improved access to highly insoluble mineral-derived products and atmospheric N₂. Our findings carry implications to sustainable reforestation and agroforestry efforts worldwide, encouraging the use and further trials of native trees simultaneously forming N₂-fixing and multiple mycorrhizal symbioses (e.g. *Acacia*, *Alnus*, *Casuarina*) either on poorly developed soils with appreciable mineral component or on highly nutrient impoverished soils in combination with rock grain application. Because mineral weathering also enables C capture through the mineral-H₂CO₃ carbonation reaction, such strategic forestry practices may not only capture C by forest re-growth but also by the means of microbe-assisted enhanced weathering.

Methods and Materials

Study sites

The three secondary forest sites in our weathering dynamics study were each 20x50m plots, located near Tarzali (-17.40, 145.61) in Queensland, Australia and their age was estimated using detailed aerial photography as previously described^{25,26}. Their soils were characterised as deeply weathered oxisols derived from basalt bedrock. The mature forest site (20x50m) in our study was part of Davies Creek National Park (-17.04, 145.61), Queensland, Australia and was characterised by deeply weathered soils derived from granite parent material (forest age was estimated using the straight line equation resulting from correlating tree DBH with age from our inventoried secondary and mature sites near Tarzali). The DBH statistics outlined in Figure 1 and in the Main Text result from an inventory tree species data from 26 secondary and 7 mature sites near Tarzali. The climatic conditions are characterised by seasonally wet mild tropical climate with rainfall in the range of 1100 to 2240 mm and temperature with a mean minimum of 10°C in winter to a mean maximum of 29°C in summer.

Nodulation estimates

Nodulation was estimated by the means of collecting 5 soil cores (cylinders with dimensions 10x10 cm) for each tree extracted 1 m away from the main tree trunk. This procedure was replicated for 8 *Acacia celsa* trees in each of the three secondary forest sites of our weathering dynamics study to a total of 24 tested trees and 120 cores. The 5 cores collected per tree were subsequently pooled and their roots separated through sieving. Nodules were carefully detached from roots using tweezers. Both collected nodules and roots were rinsed thoroughly to remove any remaining soil particles and dried for 30 min in a 70°C oven. The resulting nodule and root samples were separately weighed and nodulation was expressed as mg fresh nodule per g fresh root biomass.

Weathering dynamics

Six tree species of different families well represented in the Australian Wet Tropics flora were used to compare mineral weathering rates across the 12-48-year old forest chronosequence, namely N₂-fixing and ectomycorrhizal *Acacia celsa* (Leguminosae) and non-fixing *Alphitonia petriei* (Rhamnaceae), *Guioa lasioneura* (Sapindaceae), *Glochidion hylandii* (Phyllanthaceae), and *Neolitsea dealbata* (Lauraceae). A total of 4 polyethylene 30 µm mesh

bags (2 x basalt, 2 x dunite) containing 3 grams of crushed rock material with grain size of 250-500 μm were deposited in the immediate rhizosphere of each of the 59 tested trees (15 cm away from the main stem) using a 10x10cm soil corer to a total of 236 mineral samples. Bags were collected after 10 months and a full wet season. The contents of each of the two bags of the same rock from each tree were pooled and well-mixed to a composite rock sample for that tree (1 basalt and 1 dunite) and were subsequently analysed through portable X-ray fluorescence (XRF) Olympus Vanta M series machine set at biogeochemistry mode (the machine was rented from Sercal Non-Destructive Test Equipment Ltd., UK). To eliminate any non-scanned light elements during XRF (presented by the machine as “LE”), the amounts of scanned weathered rock from pooled basalt and dunite samples were optimised and weighed to ~0.07 g and ~0.10 g, respectively, forming a thin layer on the top of the scanning screen (window). Each sample was processed three times through XRF and averaged to minimise machine drift with its mean used for further analyses. Weathering rates were estimated by Mg loss relative to fresh material for dunite (as Mg is the main principal component) and by the widely used Product of Weathering Index (PWI)^{13,27,28} for basalt due to its more varied chemical composition. The equation for PWI is as follows:

$$PWI = mol \frac{SiO_2}{(SiO_2 + TiO_2 + Fe_2O_3 + Al_2O_3)} \times 100$$

We used all developed weathering indices and chose to use PWI on the following criteria: (1) number of chemical components used for calculation, (2) significantly lower (more weathered) values for weathered material than fresh unweathered rock, and (3) significantly correlating with dunite weathering (based on Mg loss relative to fresh material) as favourable weathering conditions for silicates are common between different rock types. Results are presented in Supplementary Table 3.

Soil analyses

Total soil chemistry was obtained through XRF of the soiled dunite-containing polyethylene mesh bags after the bags have been entirely emptied from their rock material. As the bags have spent sufficient time (10 months) in the soil environment their outer side was entirely covered in soil from the immediate vicinity. They were XRF scanned using the Olympus Vanta Alloy analysis mode. Soil for pH measurements was collected from topsoil sites (0-10cm depth) near the mineral samples. The soil pH was measured with a portable Hannah pH

meter to the second decimal space directly applying the instrument to a soil paste containing 1 g soil and 5 ml 0.1M CaCl₂ after the paste was left shaking on a horizontal shaker at 250 rpm for 1 hour.

DNA extraction and shotgun metagenome sequencing

DNA was extracted from 0.5 g of the pooled weathered rock samples that were stored at -20°C after collection in the field and prior extraction. Extraction was carried out using the MoBio PowerSoil DNA isolation kit and protocol instructions therein with slight modifications. Firstly, the amount of material added was double the recommended 0.25 g. Secondly, samples in PowerBead tubes were shaken on a horizontal shaker at 2000 rpm for 15 min to maximize extraction yield for any endolithic microorganisms. Following extraction, DNA samples were stored at -20°C before sending them for sequencing. All 34 total DNA extracts ($n=17$ dunite and $n=17$ basalt) were sequenced on one lane of the NovaSeq 150bp Paired Ends at the Edinburgh Genomics Institute, UK.

Root RNA extraction, transcriptome sequencing and de novo transcriptome assembly

Roots were collected in the field from the immediate vicinity of the weathered mesh bag samples from 10 target trees ($n=8$ *Acacia celsa* and $n=3$ *Alphitonia petriei*) using a soil corer (10x10cm) and were immediately flash frozen using liquid N₂. Upon return from the field samples were stored in -80°C before further work. Frozen root samples were handled aseptically and weighed to 250 mg input amount in a -20°C walk-in freezer room. The weighed samples were then used to extract total RNA following the RNeasy protocol for field extraction of RNA²⁹. The resulting total RNA extracts were sent in a dry shipper charged with liquid N₂ to Edinburgh Genomics Institute where they were further performed quality checks with an Agilent Bioanalyzer RNA 6000 Pico kit. Samples had exhibited significant amount of degradation likely resulting from extensive transportation with average RIN value of 2.2. Regardless of the low RIN values, we pursued further polyA tail enrichment and library construction using Illumina TruSeq stranded mRNA-seq. Samples were sequenced on the NovaSeq 150 bp Paired Ends S2 lane together with the DNA samples which were sequenced on the S1 lane. The paired-end transcriptome sequences were fed into Blast2GO Premium software and assembled using the built-in *de novo* transcriptome assembly pipeline. Assembled transcriptomes were nBLAST-ed against the NCBI database using the Cloud

Blast2GO feature to establish taxonomic identity using top hit matching. Subsequently, hits were mapped and functionally annotated using the Gene Ontology (GO) database feature within Blast2GO. We performed 2 post-assembly analyses to independently assess the integrity of the obtained data due to the low RIN values. First, we confirmed that the root samples from beneath *Acacia celsa* trees (Leguminosae) mapped to 1.5-fold more legume species than root samples from beneath *Alphitonia petriei* (Rhamnaceae) as shown in Supplementary Figure 3. Secondly, we compared nodulated ($n=3$) with non-nodulated ($n=5$) *Acacia* roots and found that previously characterised nodulation-specific pathways were significantly enriched in the nodulated roots (Supplementary Table 2) and that nodulated roots (definitely leguminous) contained significantly greater hits taxonomically mapping to legume species than non-nodulated roots (of uncertain taxonomic affinity; Supplementary Figure 3).

Shotgun metagenome assembly and an **CHANGED!**

All 34 shotgun metagenomes were processed through the MG-RAST server³⁰ using default parameters and are publicly available for download and analysis (their unique MG-RAST ID numbers and sequencing statistics can be found in Supplementary Table 4). The average size significantly differed with dunite metagenomes revealing significantly lower, 8.97 Giga base pairs (Gbp) mean, than basalt metagenomes with a mean of 11.05 Gbp (paired two-tailed t-test, $P<0.01$). Annotation used in our analyses were against the RefSeq protein database and hierarchical clustering was achieved via the Subsystem database classifying functional reads into levelled pathways (Level 3, 2, 1). Comma-delimited table files at e-5, 40% similarity for all levels and functional reads were downloaded for further analyses. **Further information on sequencing statistics can be found in Appendix 2.**

Correlation clustering analyses

Correlation clustering was performed on Level 3 pathways with weathering rates inserted into the input table so that they cluster with their strongly correlating pathways and with pathways that strongly correlate with pathways strongly correlating with weathering revealing the metabolic conditions favouring or co-varying with enhanced weathering. Consequently, we constructed Spearman correlation test correlation matrix containing rho values. The resulting matrix and its constituents were clustered using complete Manhattan distance and the heatmap.2 function in the R environment.

Bradyrhizobial nifK and fungal taxonomic analyses

Fe-Mo nitrogenase genes can be used to dissect the diversity of both possible free-living and nodulating diazotrophs. Similarly, the large ribosomal subunit (LSU) genes can be used for taxonomic analyses of eukaryotes. We downloaded the FASTA files for all the Fe-Mo nitrogenase beta subunit (*nifK*) annotated hits and the eukaryotic and archaeal LSU gene (Level 3 hierarchical cluster) hits from each of our metagenomes using MG-RAST interface. The FASTA files were fed into Blast2GO and nBLAST was performed against the NCBI database using the in-built Cloud computing. Next, we used the best matching hit (according to its e-value) for each of the sequences to assemble a species table for all hits for each sample. Species were clustered into genera and samples were merged together into one taxonomic analysis file. To distinguish ectomycorrhizal fungi within the list of eukaryotic and archaeal LSU genera, we used the detailed tabulated list of documented ectomycorrhizal fungal genera assembled by Tedersoo and Smith³¹. The sum of ectomycorrhizal fungi (as a proportion of all eukaryotic and archaeal LSU reads) were multiplied by the abundance of the eukaryotic and archaeal LSU gene within each metagenome to obtain a surrogate measure for the abundance of ectomycorrhizal fungi within each metagenomic microbial community. Similarly, *nifK* sequences that best matched to *Bradyrhizobium nifK* sequences were summed and then multiplied by the *nifK* gene abundance in their representative metagenome. The abundance of bradyrhizobial and ectomycorrhizal fungi within the root transcriptomes were derived by a similar approach based on nBLAST of all transcripts in the Blast2GO environment and subsequent use of the best matching taxonomic hits.

Statistical analyses

All statistical analyses and their P value and correlation r and rho results are presented in the Main Text and/or in the Figures and Figure Text. Generally, normally distributed data were tested with parametric tests, whereas data lacking normal distribution were tested with their non-parametric test equivalents. Binning into low and high P leaching categories was based on the median P leaching rate due to significant outliers preventing reliable correlational analyses. Due to their highly variable yields (several hundred transcripts to several thousand transcripts), root transcriptomic GO pathways correlations with basalt P leaching rates were achieved by excluding any transcriptomes that lacked the given GO pathway due to shallow sequencing – as such GO pathways were correlated based on 3, 4, 5, 6 or 7 entries with

different cut-off Pearson r values – essentially using GO term-specific rarefaction ($n \geq 3$). Only transcriptomes from roots beneath *Acacia* were correlated with P leaching to avoid species effects. One transcriptome was excluded due to a lacking P leaching rate ($n=8-1=7$).

References

1. Batterman, S.A., Hedin, L.O., Van Breugel, M., Ransijn, J., Craven, D.J. and Hall, J.S., 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature*, 502(7470), p.224.
2. Barron, A.R., Purves, D.W. and Hedin, L.O., 2011. Facultative nitrogen fixation by canopy legumes in a lowland tropical forest. *Oecologia*, 165(2), pp.511-520.
3. Vitousek, P.M., Menge, D.N., Reed, S.C. and Cleveland, C.C., 2013. Biological nitrogen fixation: rates, patterns and ecological controls in terrestrial ecosystems. *Philosophical Transactions of the Royal Society B: Biological Sciences*, 368(1621), p.20130119.
4. Yeo, W.L.J. and Fensham, R.J., 2014. Will *Acacia* secondary forest become rainforest in the Australian Wet Tropics?. *Forest ecology and management*, 331, pp.208-217.
5. Houghton, R.A. and Nassikas, A.A., 2018. Negative emissions from stopping deforestation and forest degradation, globally. *Global change biology*, 24(1), pp.350-359.
6. Taubert, F., Fischer, R., Groeneveld, J., Lehmann, S., Müller, M.S., Rödiger, E., Wiegand, T. and Huth, A., 2018. Global patterns of tropical forest fragmentation. *Nature*, 554(7693), p.519.
7. Ernst, C., Mayaux, P., Verhegghen, A., Bodart, C., Christophe, M. and Defourny, P., 2013. National forest cover change in Congo Basin: deforestation, reforestation, degradation and regeneration for the years 1990, 2000 and 2005. *Global change biology*, 19(4), pp.1173-1187.
8. Davidson, E.A., de Carvalho, C.J.R., Figueira, A.M., Ishida, F.Y., Ometto, J.P.H., Nardoto, G.B., Sabá, R.T., Hayashi, S.N., Leal, E.C., Vieira, I.C.G. and Martinelli, L.A., 2007. Recuperation of nitrogen cycling in Amazonian forests following agricultural abandonment. *Nature*, 447(7147), p.995.
9. Wurzbarger, N. and Hedin, L.O., 2016. Taxonomic identity determines N₂ fixation by canopy trees across lowland tropical forests. *Ecology letters*, 19(1), pp.62-70.
10. Hopkins, M.S., Reddell, P., Hewett, R.K. and Graham, A.W., 1996. Comparison of root and mycorrhizal characteristics in primary and secondary rainforest on a metamorphic soil in North Queensland, Australia. *Journal of Tropical Ecology*, 12(6), pp.871-885.
11. Newbery, D.M., Alexander, I.J. and Rother, J.A., 1997. Phosphorus dynamics in a lowland African rainforest: the influence of ectomycorrhizal trees. *Ecological Monographs*, 67(3), pp.367-409.
12. Hart, T.B., Hart, J.A. and Murphy, P.G., 1989. Monodominant and species-rich forests of the humid tropics: causes for their co-occurrence. *The American Naturalist*, 133(5), pp.613-633.
13. Fiantis, D., Nelson, M., Shamshuddin, J., Goh, T.B. and Van Ranst, E., 2010. Determination of the geochemical weathering indices and trace elements content of new volcanic ash deposits from Mt. Talang (West Sumatra) Indonesia. *Eurasian Soil Science*, 43(13), pp.1477-1485.
14. Andrews, M. and Andrews, M.E., 2017. Specificity in legume-rhizobia symbioses. *International Journal of Molecular Sciences*, 18(4), p.705.
15. Yeoh, Y.K., Paungfoo-Lonhienne, C., Dennis, P.G., Robinson, N., Ragan, M.A., Schmidt, S. and Hugenholtz, P., 2016. The core root microbiome of sugarcane cultivated under varying nitrogen fertilizer application. *Environmental microbiology*, 18(5), pp.1338-1351.
16. Tang, C., Unkovich, M.J. and Bowden, J.W., 1999. Factors affecting soil acidification under legumes. III. Acid production by N₂-fixing legumes as influenced by nitrate supply. *The New Phytologist*, 143(3), pp.513-521.
17. Schwalbach, M.S., Tripp, H.J., Steindler, L., Smith, D.P. and Giovannoni, S.J., 2010. The presence of the glycolysis operon in SAR11 genomes is positively correlated with ocean productivity. *Environmental microbiology*, 12(2), pp.490-500.
18. Lin, T.F., Huang, H.I., Shen, F.T. and Young, C.C., 2006. The protons of gluconic acid are the major factor responsible for the dissolution of tricalcium phosphate by *Burkholderia cepacia* CC-A174. *Bioresource Technology*, 97(7), pp.957-960.
19. Duff, R.B., Webley, D.M. and Scott, R.O., 1963. Solubilization of minerals and related materials by 2-ketogluconic acid-producing bacteria. *Soil Science*, 95(2), pp.105-114.

21. Leslie, K., Sturm, A., Stotler, R., Oates, C.J., Kyser, T.K. and Fowle, D.A., 2015. *Marinobacter* bacteria associated with a massive sulphide ore deposit affect metal mobility in the deep subsurface. *Geochemistry: Exploration, Environment, Analysis*, 15(4), pp.319-326.
22. Bennett, P.C., Rogers, J.R., Choi, W.J. and Hiebert, F.K., 2001. Silicates, silicate weathering, and microbial ecology. *Geomicrobiology Journal*, 18(1), pp.3-19.
23. Hamilton, W.A., 1985. Sulphate-reducing bacteria and anaerobic corrosion. *Annual review of microbiology*, 39(1), pp.195-217.
24. Epihov, D.Z., Batterman, S.A., Hedin, L.O., Leake, J.R., Smith, L.M. and Beerling, D.J., 2017. N₂-fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?. *Proceedings of the Royal Society B: Biological Sciences*, 284(1860), p.20170370.
25. Goosem, M., Paz, C., Fensham, R., Preece, N., Goosem, S. and Laurance, S.G., 2016. Forest age and isolation affect the rate of recovery of plant species diversity and community composition in secondary rain forests in tropical Australia. *Journal of Vegetation Science*, 27(3), pp.504-514.
26. Paz, C.P., Goosem, M., Bird, M., Preece, N., Goosem, S., Fensham, R. and Laurance, S., 2016. Soil types influence predictions of soil carbon stock recovery in tropical secondary forests. *Forest Ecology and Management*, 376, pp.74-83.
27. Sour, B., Watanabe, M. and Sakagami, K., 2006. Contribution of Parker and Product indexes to evaluate weathering condition of Yellow Brown Forest soils in Japan. *Geoderma*, 130(3-4), pp.346-355.
28. Reiche, P., 1950. *A survey of weathering processes and products* (No. 3). University of New Mexico Press.
29. Breitler, J.C., Campa, C., Georget, F., Bertrand, B. and Etienne, H., 2016. A single-step method for RNA isolation from tropical crops in the field. *Scientific reports*, 6, p.38368.
30. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A. and Wilkening, J., 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9(1), p.386.
31. Tedersoo, L. and Smith, M.E., 2013. Lineages of ectomycorrhizal fungi revisited: foraging strategies and novel lineages revealed by sequences from belowground. *Fungal biology reviews*, 27(3-4), pp.83-99.
32. Dinnage, R., Simonsen, A.K., Cardillo, M., Thrall, P.H., Barrett, L.G. and Prober, S.M., 2018. Larger legume plants host a greater diversity of symbiotic nitrogen-fixing bacteria. *bioRxiv*, p.246611.
33. Wolińska, A., Kuźniar, A., Zielenkiewicz, U., Banach, A., Izak, D., Stępniewska, Z. and Błaszczuk, M., 2017. Metagenomic analysis of some potential nitrogen-fixing bacteria in arable soils at different formation processes. *Microbial ecology*, 73(1), pp.162-176.

Figures and Figure Text

(Epihov *et al.*, 2018)

Figure 1. Diameter at breast height (DBH) data from >2,700 stems from 26 Australian wet tropical secondary forest sites of different post-abandonment ages reveal that the early pioneering N₂-fixing *A. celsa* grows faster than all non-fixers combined and our target non-fixing trees: a. *Acacia celsa* (Leguminosae); b. all non-fixers; c. *Alphitonia petriei* (Rhamnaceae); d. *Guoia lasioneura* (Sapindaceae); e. *Glochidion hylandii* (Phyllanthaceae); f. *Neolitsea dealbata* (Lauraceae);

Table 1. Patterns of abundance in *A. celsa* suggest that monodominance (based on % stem basal area>30%) occurs in later stages of secondary succession; N₂-fixers in mature forests are not *A. celsa* but *Archidendron* and *Austrosteenisia* spp. R² values are results from Pearson correlation test.

Figure 2. Dunite and basalt rock weathering rates are higher during early successional stages but decrease in later successional and mature forests with the successional monodominant N₂-fixer *Acacia celsa* exhibiting higher leaching of the macroelements P and K from basalt than non-fixers in secondary forests. **a.** Dunite Mg declines in weathered minerals with largest declines observed in early successional forests (ANOVA, $P < 0.001$); **b.** Basalt Product of Weathering Index (PWI) declines relative to fresh basalt signifying greater weathering rates in young forests and lower ones in late successional and mature (120-year old) forests (ANOVA, $P < 0.001$); **c.** P leaching patterns across the chronosequence shows consistently greater P leaching in the rhizosphere of *A. celsa* (Kruskal-Wallis test, $P < 0.05$); **d.** *A. celsa* rhizospheres reveal significantly greater P and K declines from basalt rock material relative to non-fixing trees (Mann-Whitney test). Fisher's LSD Post-Hoc test was applied to determine the significant differences between groups.

Figure 3. Nitrogen-driven feedbacks resulting from increased inputs of symbiotically fixed N mediate metagenomic increases in acidifying and lithotrophic metabolic pathways clustering with enhanced weathering rates. **a.** The highest weathering rates observed in 20-year old forests co-occur with peaks in symbiotic N₂-fixation (nodulation) in *Acacia celsa* trees and the relative abundance of genes assigned to ammonia monooxygenase (catalysing the rate-limiting step in nitrification: $\text{NH}_3 \rightarrow \text{NO}_2^-$) and N₂-fixation in the rock metagenomes suggesting tight link between N cycling and weathering dynamics; **b.** Spearman correlation rho heat-map matrices for abundance of metabolic pathways ($n=1130$) in basalt and dunite metagenomes ($n=17$ each) input together with weathering rate reveal three clusters (*sensu stricto* – 1&2 and *sensu lato* – 3) of pathways positively associated with weathering rates and one (4) of pathways broadly negatively associated with weathering rates. Clustering is based on complete Manhattan clustering method; **c.** Mean Spearman rho correlation values for each of the four identified clusters with rock weathering rates in comparison to the overall average for all pathways confirming that clusters 1&2 contain metabolic pathways positively associated with weathering rates (ANOVA, $P < 0.001$ for both minerals; Post-Hoc Tukey HSD test); **d.** Proposed N-driven feedbacks affecting the function of rock-associated microbial communities using metabolic pathways from clusters 1&2 in both rock materials and their link to enhanced weathering through organic acids and H⁺ generation (red typeface). Pathways and/or genes in bold are of key importance to weathering and are discussed in the Main Text.

Figure 4. *Acacia celsa* recruits multiple symbionts leading to improved access to mineral nutrients. **a.** Root transcriptomics reveal that nodulated *Acacia* roots have significantly higher *Bradyrhizobium* transcripts (ANOVA, $P < 0.01$, Fisher's LSD Post-Hoc test) suggesting that *Bradyrhizobium* spp. are the likely nodulating agents of *A. celsa*; **b.** Basalt metagenomics show significantly higher bradyrhizobial *nifK* (Fe-Mo nitrogenase beta subunit) gene abundance in *A. celsa* relative to the pioneering non-fixer *Alphitonia petriei* suggesting that symbiotic attraction may affect the microbial community of basalt (two-tailed t-test, $P < 0.05$); **c.** Root transcriptomes reveal greater abundance of ectomycorrhizal (EM) fungal transcripts in *Acacia* roots compared to the roots of non-fixer *A. petriei* (Mann-Whitney test, $P < 0.05$); **d.** Paired comparison of metagenomics eukaryotic and archaeal LSU reveals significantly higher EM fungal abundance in basalt than dunite (Wilcoxon's paired ranked test, $P < 0.05$); **e.** Analysis of the ribosomal large subunit genes (LSUs) of eukaryotes and archaea in basalt metagenomes indicate significantly greater abundance of EM fungi and overall cumulative abundance of mycorrhizal fungi (EM + arbuscular mycorrhizal (AM) fungi) in basalt from *A. celsa* rhizosphere than that beneath the non-fixer *A. petriei* (two-way ANOVA, fixer/non-fixer $P < 0.05$, Sidak's multiple comparison test); **f.** Successional dynamics in *nifK*, EM and AM LSUs from basalt reveal peaks in *nifK* matching that of nodulation (see Figure 3a) and peak in EM LSUs in the oldest 48-year forests with the highest P leaching (see Figure 2c); **g.** Mycorrhizal community profiling shows multiple EM and AM symbiotic recruits with specificities in roots, basalt and dunite rock

material; **h.** Greater abundance of *Pisolithus* and *Hebeloma* symbionts within the basalt EM community is linked to high P leaching (Mann-Whitney test, $P < 0.01$ and $P < 0.05$).

Table 2. Root transcriptomics of *Acacia celsa* roots elucidate strong responses to basalt P leaching

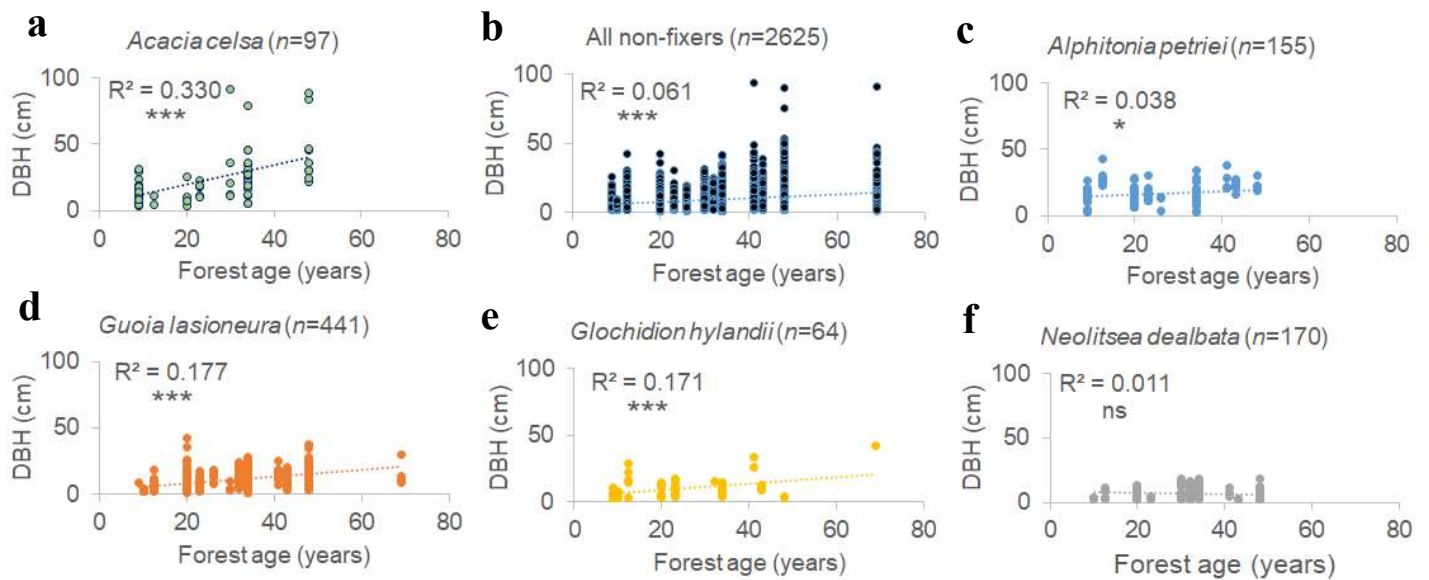


Figure 1. (Epihov *et al.*, 2018.)

Table 1. (Epihov *et al.*, 2018.)

N ₂ -fixers	Forest age (n = number of 500 m ² plots)				
	12.5 (n=3)	20 (n=4)	23-32 (n=4)	34-48 (n=10)	Mature (n=7)
Maximum BA%	1.36	12.35	57.16	64.18	2.25
Mean BA%	0.45	3.09	30.13	16.46	0.35
Monodominance (>30% BA)	No	No	Yes (2/4)	Yes (4/10)	No

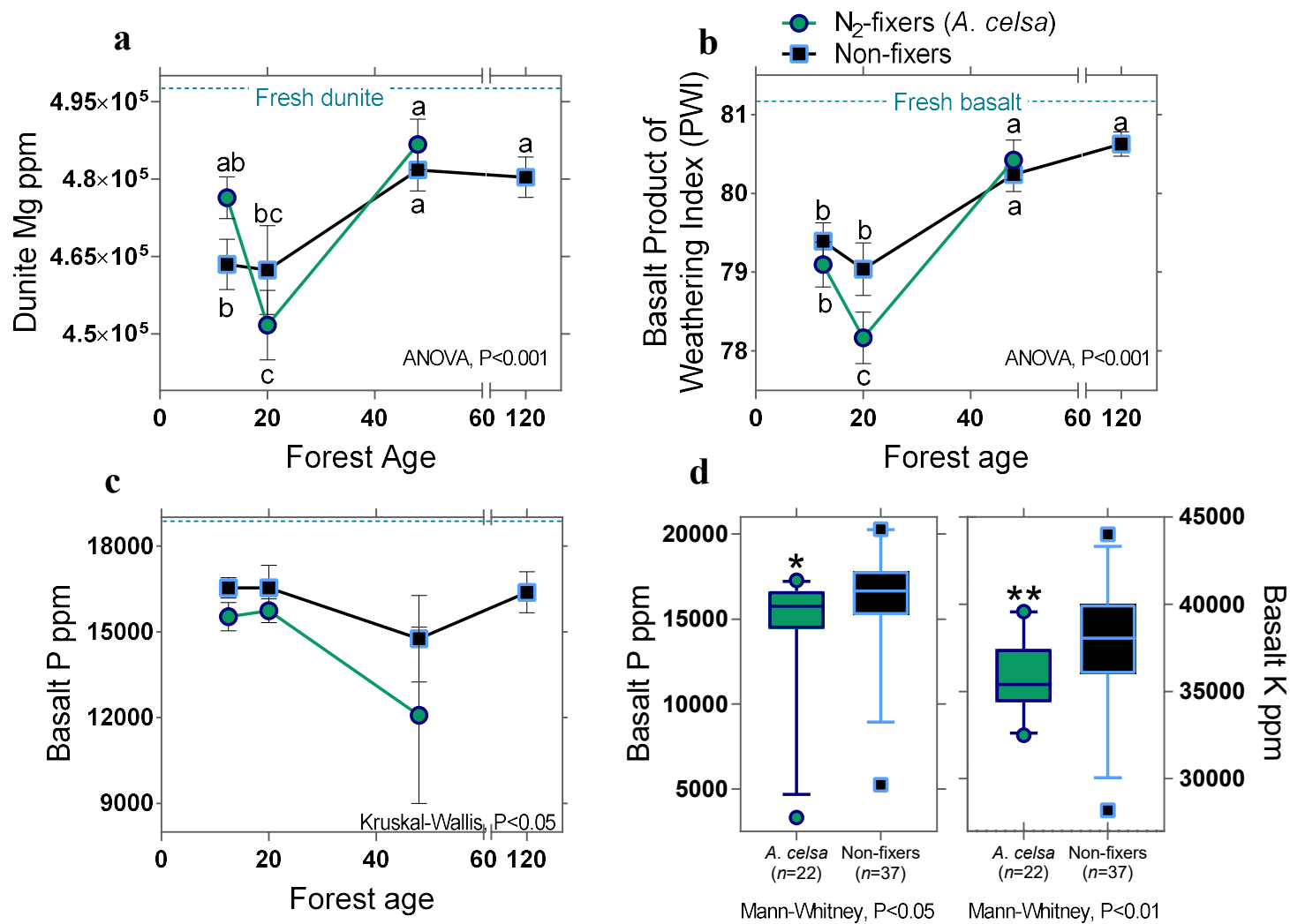
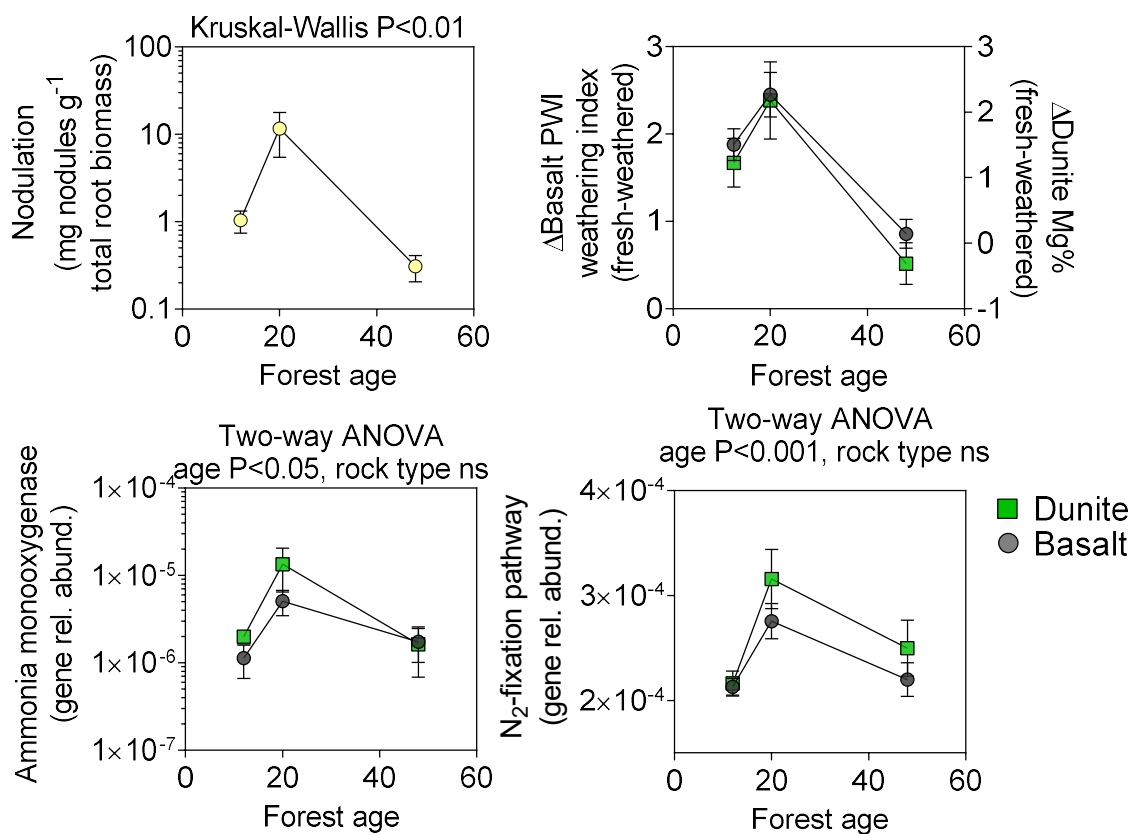
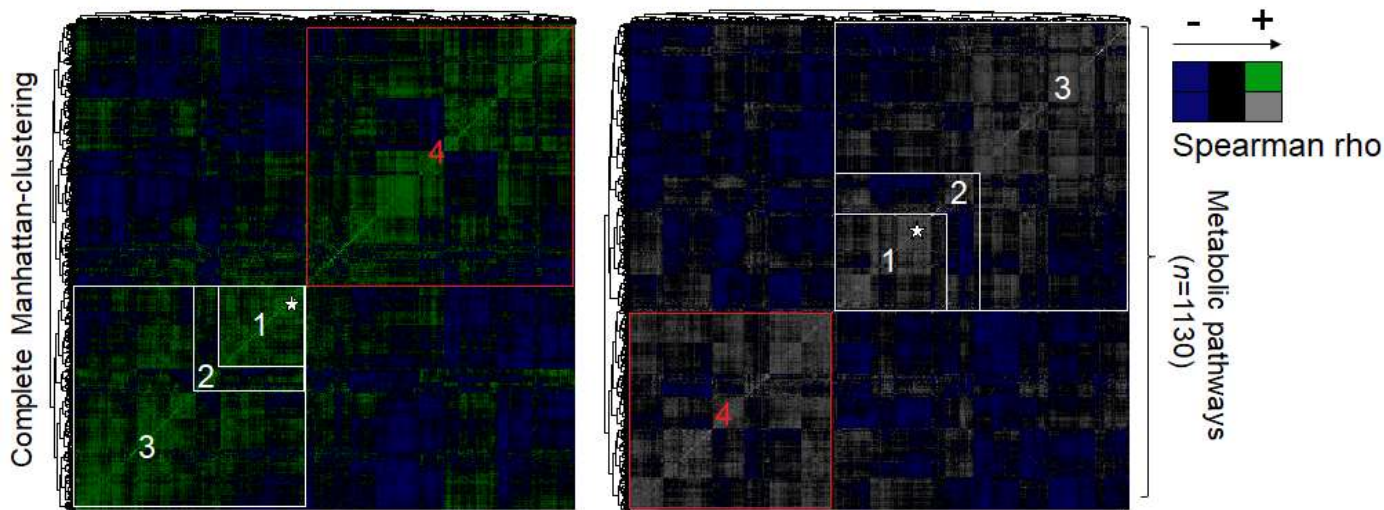
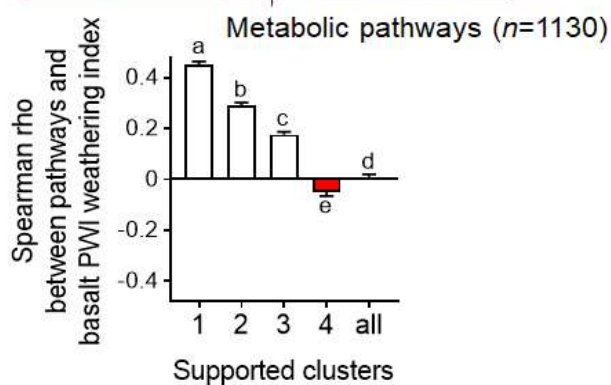
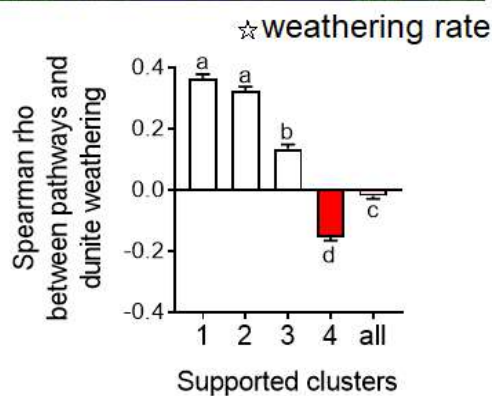


Figure 2. (Epihov *et al.*, 2018.)

a**b**

Correlation heat-map matrices

Dunite metagenomes ($n=17$)Basalt metagenomes ($n=17$)**c**

d

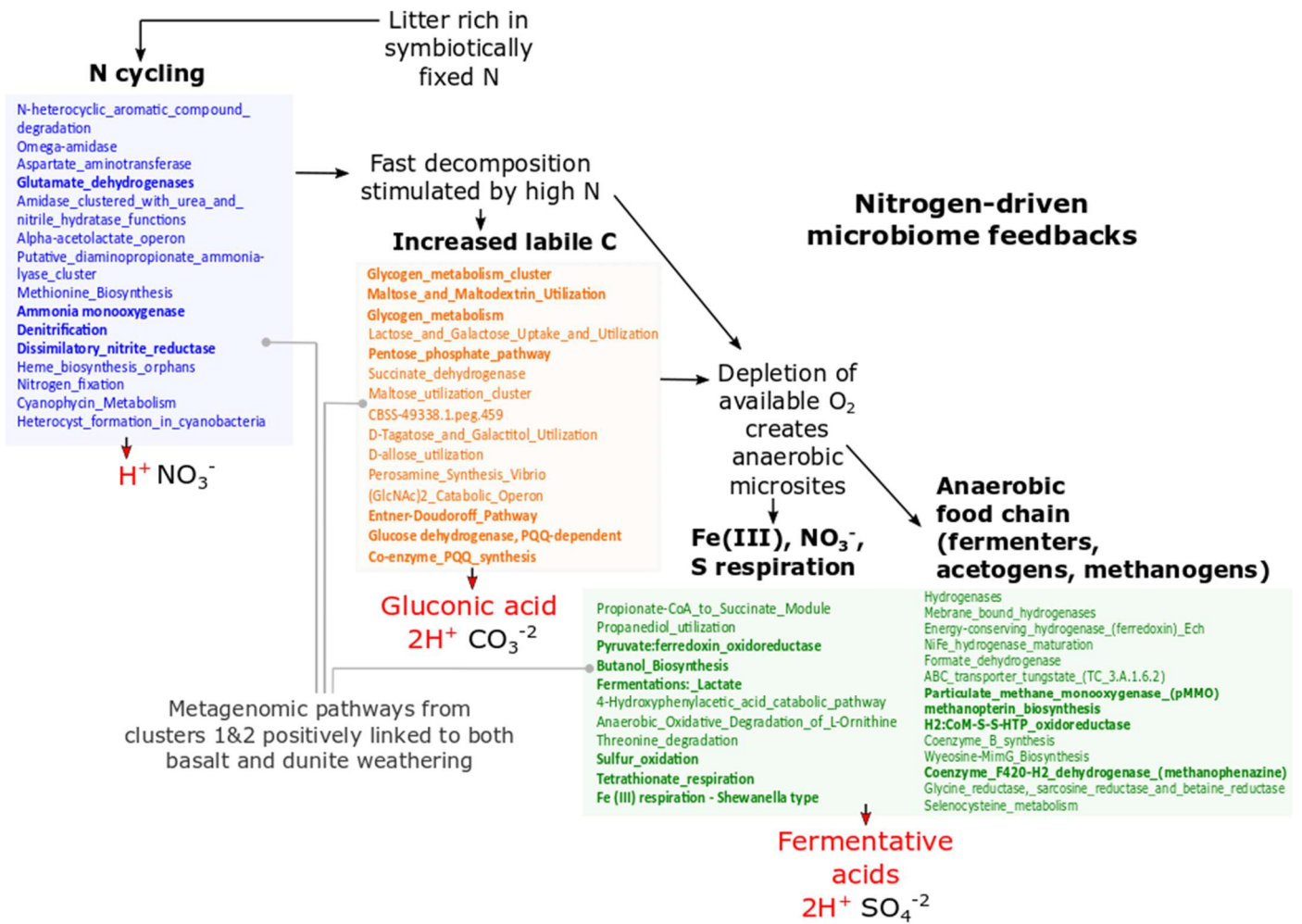


Figure 3. (Epihov *et al.*, 2018.)

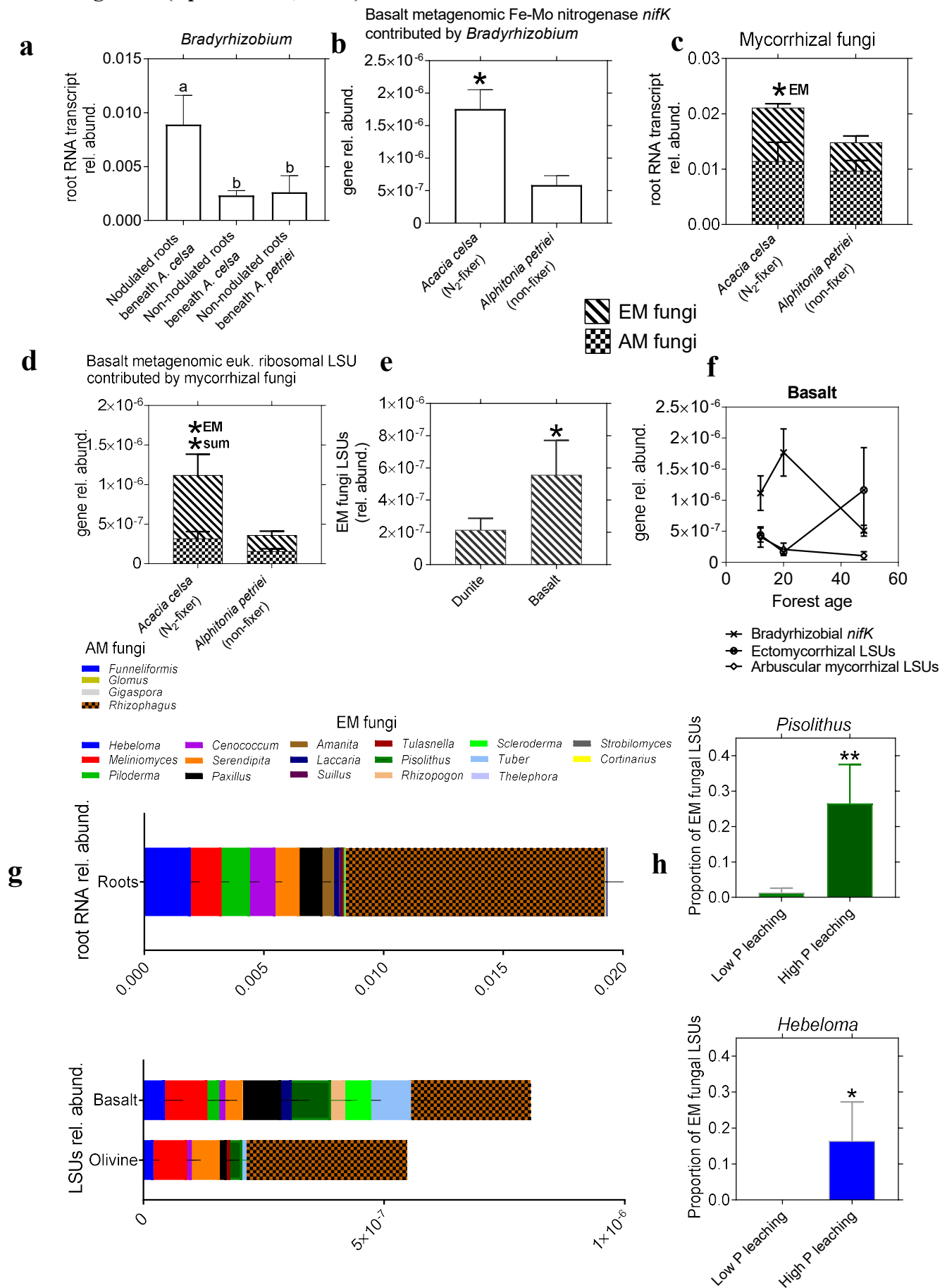


Figure 4. (Epihov *et al.*, 2018.)

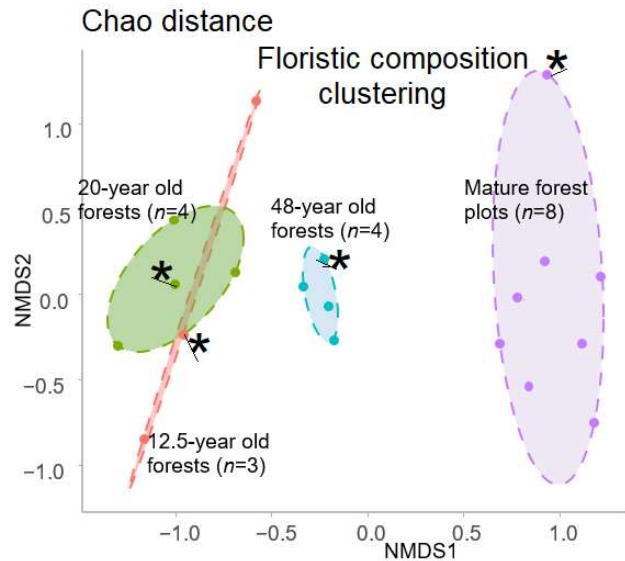
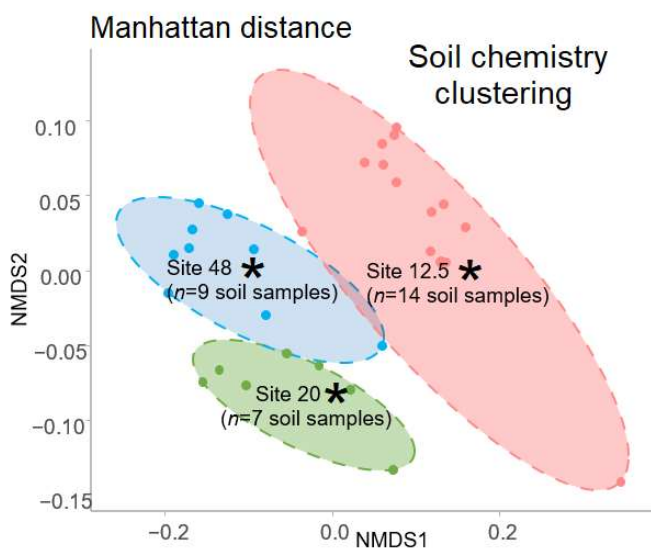
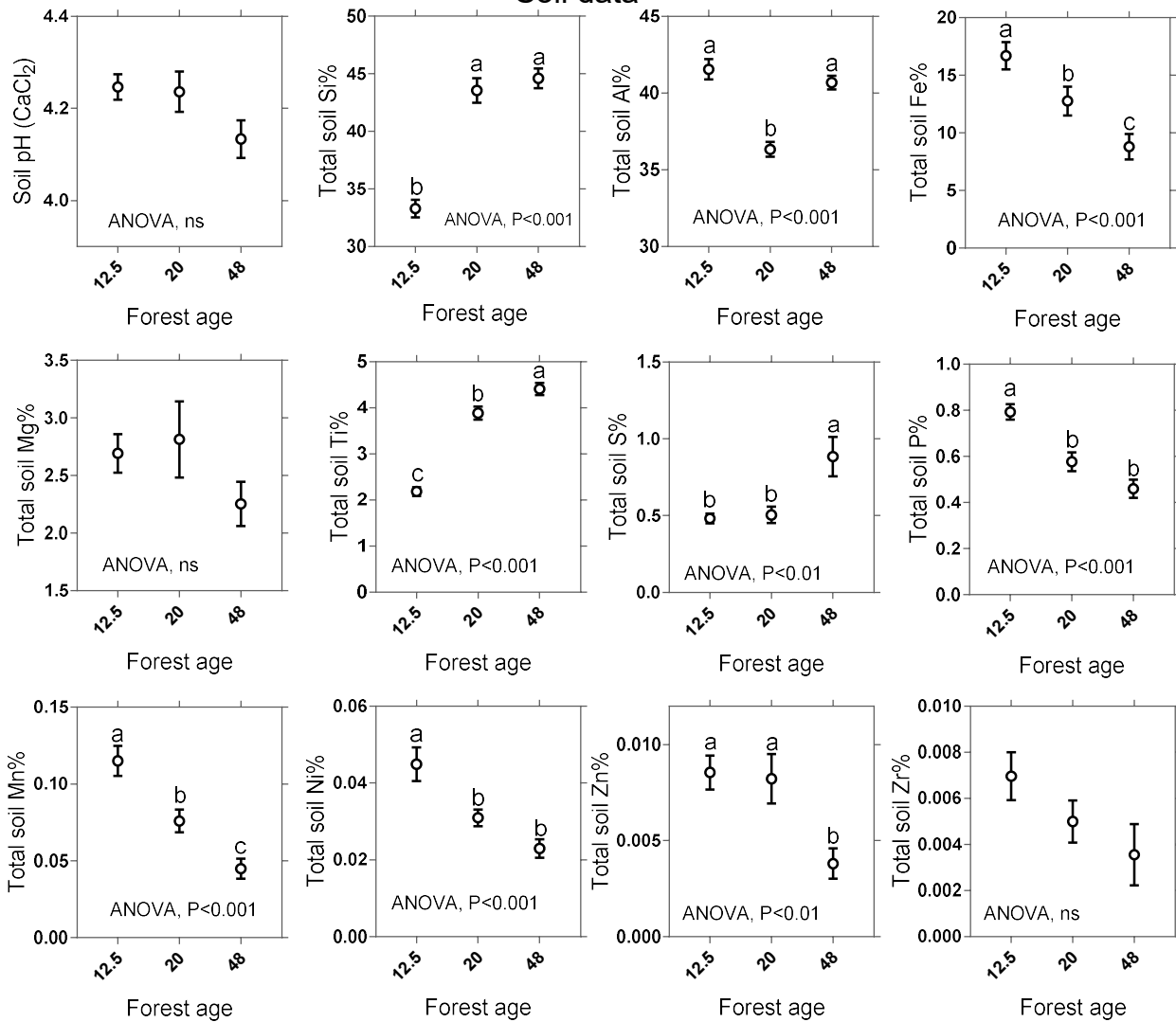
Table 2. (Epihov *et al.*, 2018.)

Positive correlations between abundance of GO pathways in root transcriptomes and basalt P leaching					
GO term	Pearson test P-value	No. of transcriptomes in which GO term present	Stimulating P leaching	Stimulated by P leaching	Mechanism
Proton transmembrane transport	0.022†	7	yes		proton extrusion enables P dissolution
Oxidation reduction process	0.058†	7	yes		up-regulation of energy metabolism may generate more CO ₂ , organic acids, protons etc.
Aerobic electron transport chain	0.011	3	yes		up-regulation of energy metabolism may generate more CO ₂ , organic acids, protons etc.
ATP synthesis coupled electron transport	0.0009	7	yes	yes	up-regulation of energy metabolism may generate more CO ₂ , organic acids, protons etc.
Signal transduction by protein phosphorylation	0.063	6		yes	high P availability may promote this P-dependent process
Phosphorelay signal transduction system	0.034	6		yes	high P availability may promote this P-dependent process
Protein phosphorylation	0.010	7		yes	high P availability may promote this P-dependent process
Carbohydrate phosphorylation	0.006	5		yes	high P availability may promote this P-dependent process
Phospholipid translocation	0.031	4		yes	high P availability may promote this P-dependent process
Defense response	0.011	5		yes	high P leaching may trigger competitive and symbiotic interactions between plant roots and microorganisms in the rhizosphere
Response to biotic stimulus	0.007	5		yes	high P leaching may trigger competitive and symbiotic interactions between plant roots and microorganisms in the rhizosphere

†Values excluding a single very high P leaching outlier

Supplementary Figures and Text

Soil data



*forest plots (sites) used in our weathering study

Supplementary Figure 1. Soil chemistry data and floristic composition suggest that the latter (typically a product of successional dynamics) is the more likely driver of weathering differences than the former with sites of high weathering rates (our 12.5 site and 20-year old site) exhibiting overlapping floristic composition distinct from that of 48-year old forests (our 48-year old site revealed the lowest weathering rates). These findings support the view that differences in weathering between the sites used in our weathering studies (marked with “*”) are more likely age/successional effects rather than isolated site effects.



Supplementary Figure 2. Nodulation patterns in legumes of Australian tropical forests: **a.** highly nodulated roots from a 20-year old *Acacia celsa* tree in a secondary forest; **b.** nodulated aerial adventitious roots in a 48-year old *Acacia celsa* tree in a secondary forests; **c.** nodulation in a *Archidendron* sp. in a mature forest setting with its odd-pinnate leaf photographed in the upper right corner; **d.** nodules (not photographed) were also found on the underground roots of this legume liana found in mature forests. The legume liana was professionally ID-ed (Steve McKenna, botanist at Department of Agriculture and Water Resources, Queensland) as *Austrosteenisia stipularis*.

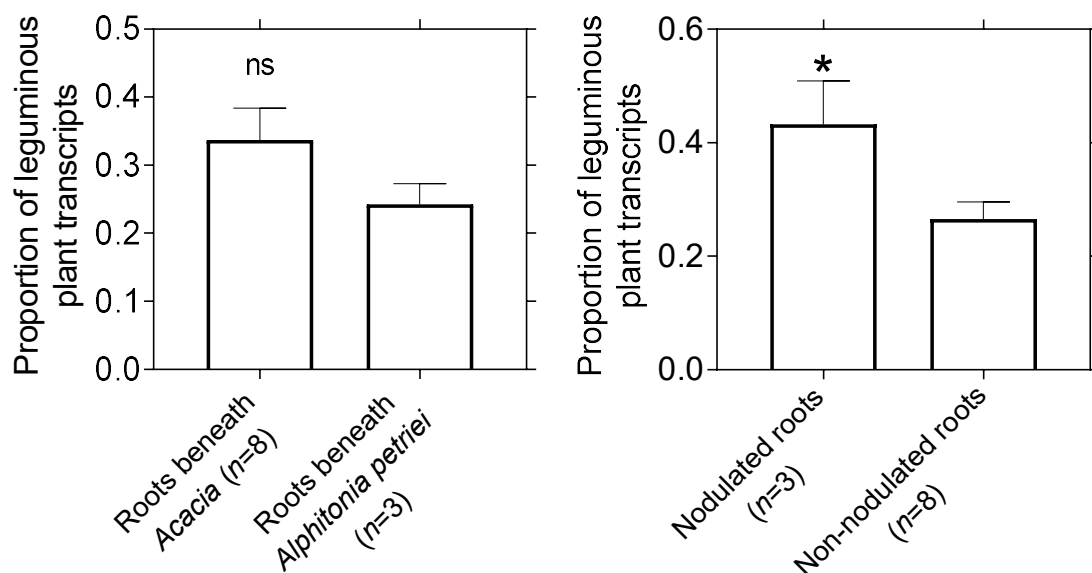
Supplementary Table 1. Significant positive correlations between basalt metagenomic carbonic anhydrase gene abundance and Level 3 Subsystem pathways involved in respiration, C cycling and N conversions.

Metagenomic pathway	Spearman rho	
	(Carbonic anhydrase)	P-value
Carbonic anhydrase	1.000	***
Respiratory_dehydrogenases_1	0.794	***
Pyrroloquinoline_Quinone_biosynthesis	0.691	**
Gluconate 2-dehydrogenase (EC 1.1.99.3), membrane-bound, flavoprotein	0.612	*
Nitrogen_fixation	0.665	*
Glutamine,_Glutamate,_Aspartate_and_Asparagine_Biosynthesis	0.591	*
Glutamate_and_Aspartate_uptake_in_Bacteria	0.548	*
Denitrification	0.546	*
Aromatic_amino_acid_degradation	0.685	**
Tryptophan_catabolism	0.735	**
Nitrosative_stress	0.737	**
Nitrate_and_nitrite_ammonification	0.484	^

Spearman correlation tests, ***P<0.001, **P<0.01, *P<0.05, ^P<0.10.

Supplementary Table 2. Validation of our *de novo* assembled root transcriptomes of *Acacia celsa* revealing enrichment in key GO pathways with known function in rhizobial infection and nodulation

Up-regulated GO pathways in nodulated roots				
GO term	Mann-Whitney test P-value	Non-nodulated roots	Nodulated roots	Role in nodulation
Oxygen transport	0.010	0	0.0035	To limit O ₂ concentration allowing properly functioning nitrogenases ¹
Spermidine biosynthetic process	0.010	0	0.0032	Bioavailable form of fixed N ²
Arginine catabolic process	0.010	0	0.0031	Bioavailable form of fixed N ³
Glutamine biosynthetic process	0.081	0.0010	0.0034	Bioavailable form of fixed N ⁴
Translation	0.053	0.0253	0.0408	High protein synthesis in stage III and IV of the rhizobial infection ⁵
Translational elongation	0.010	0	0.0075	High protein synthesis (as above)
Transcription, RNA templated	0.039	0.0005	0.0084	High protein synthesis (as above)
RNA-dependent DNA biosynthetic process	0.025	0.0071	0.01621	Division of cells (requiring DNA replication) within the symbiosomes or lateral meristem activity within the roots?
Aerobic electron transport chain	0.081	0.0014	0.0106	Up-regulated energy metabolism due to high N?
Glutathione metabolic process	0.010	0	0.0031	Anti-oxidant function important for nodulation by <i>Bradyrhizobium</i> ⁶
S-adenosylmethionine biosynthetic process	0.010	0	0.0035	Metabolic product in nodules ⁷



Supplementary Figure 3. Validation of root transcriptome analyses revealing more nBLAST hits mapping to leguminous transcripts, particularly in transcriptomes of nodulated roots.

Three of the root samples from beneath *Acacia* contained nodulated roots, with the remaining 5 samples from beneath *Acacia* showing no nodules. The non-nodulated root group includes the latter 5 and those from beneath *A. petriei*. Two-tailed t-test, * $P < 0.05$, ns $P > 0.05$. nBLAST hits were obtained against the NCBI database through Blast2GO.

Supplementary Table 3. Basalt weathering indices and related metrics.

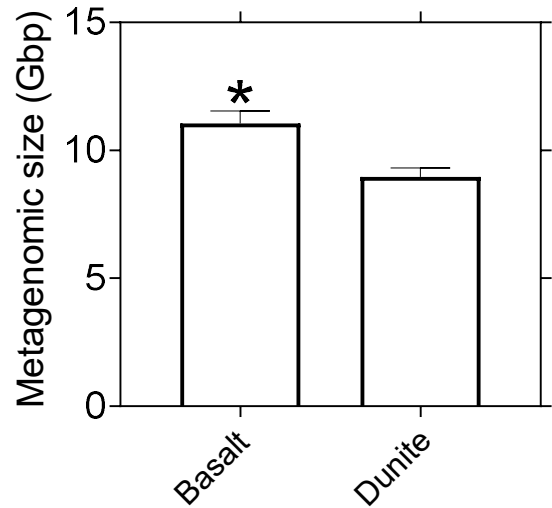
Weathering index	Fresh basalt	Weathered basalt: Non-fixers	Weathered basalt: N ₂ -fixers	Correct pattern	Test	P-value	Significance	Fisher's LSD test	Correlation with [Mg] olivine
Ruxton's index (R)	6.735	6.114	5.967	Yes	ANOVA	$P < 0.05$	Yes	a/b/b	Yes $R = 0.47^{***}$
Product of weathering index (PWI) ^s	81.158	79.499	79.061	Yes	ANOVA	$P < 0.01$	Yes	a/b/b	Yes $R = 0.35^{**}$
Weathering index of Parker (WIP)	73.898	71.459	71.372	Yes	ANOVA	$P > 0.10$	No	NA	No
Weathering Potential Index (WPI)	12.263	11.226	11.708	Yes	ANOVA	$P > 0.10$	No	NA	No
Vogt's Index	1.182	1.453	1.392	Yes	Kruskal-Wallis	$0.05 < P < 0.1$	No	NA	No
Chemical Index of Alteration (CIA)	60.537	61.947	62.141	Yes	Kruskal-Wallis	$P > 0.10$	No	NA	No
Chemical Index of Weathering (CIW)	67.060	68.526	68.290	Yes	Kruskal-Wallis	$P > 0.10$	No	NA	No
Plagioclase Index of Alteration (PIA)	63.084	64.776	64.818	Yes	Kruskal-Wallis	$P > 0.10$	No	NA	No
Silica-Titania Index (STI)	73.980	73.184	72.937	Yes	Kruskal-Wallis	$P > 0.10$	No	NA	No

^sPWI was selected as the main measure of basalt weathering due to: (1) its sensitivity in comparing fresh to weathered material, (2) its correlation with [Mg] dunite weathering signifying its sensitivity to environmental conditions favouring silicate weathering (regardless of rock type), and (3) its derivation as it is based on Si, Ti, Fe and Al measurements, whereas Ruxton's index (R) is only using basalt Si and Al. The

exact equations for all weathering indices can be found summarised in ⁸and references therein.

Supplementary Table 4. Sequencing statistics for all soil-mineral shotgun metagenomes

MG-RAST ID	Sample	Metagenome size (Gbp)	No. of sequences (x10 ⁶)	Rock type
mgm4800924.3	58	8.84	57.52	basalt
mgm4801675.3	59	13.05	84.79	basalt
mgm4801671.3	60	14.53	94.80	basalt
mgm4801670.3	61	12.47	81.25	basalt
mgm4801667.3	62	10.66	69.40	basalt
mgm4801673.3	63	11.77	76.52	basalt
mgm4801674.3	64	13.49	88.35	basalt
mgm4801669.3	65	10.53	68.72	basalt
mgm4801668.3	66	9.82	63.97	basalt
mgm4803487.3	75	9.96	65.12	basalt
mgm4802743.3	76	13.97	91.22	basalt
mgm4802748.3	78	7.84	50.85	basalt
mgm4802741.3	79	9.04	58.59	basalt
mgm4803662.3	80	10.70	68.89	basalt
mgm4803664.3	81	12.95	83.95	basalt
mgm4803665.3	82	9.23	59.75	basalt
mgm4800921.3	49	9.84	64.24	dunite
mgm4800925.3	50	9.71	63.17	dunite
mgm4800917.3	51	7.42	48.14	dunite
mgm4800923.3	52	11.45	74.59	dunite
mgm4800920.3	53	10.57	68.78	dunite
mgm4800922.3	54	8.43	55.01	dunite
mgm4800926.3	55	8.77	57.07	dunite
mgm4800918.3	56	10.40	67.90	dunite
mgm4800919.3	57	11.69	76.22	dunite
mgm4801672.3	67	7.64	49.58	dunite
mgm4801676.3	68	8.13	52.94	dunite
mgm4802747.3	69	8.72	56.54	dunite
mgm4802745.3	70	7.76	50.27	dunite
mgm4802744.3	71	9.35	60.23	dunite
mgm4802746.3	72	7.86	51.02	dunite
mgm4802742.3	73	7.33	47.73	dunite
mgm4802740.3	74	7.41	48.21	dunite



Supplementary Figure 4. Basalt supports significantly higher metagenomic sizes despite the same extraction and prep-up procedures suggesting higher microbial biomass.

Supplementary References

1. Appleby, C.A., 1984. Leghemoglobin and *Rhizobium* respiration. *Annual Review of Plant Physiology*, 35(1), pp.443-478.
2. Efrose, R.C., Fletmetakis, E., Sfichi, L., Stedel, C., Kouri, E.D., Udvardi, M.K., Kotzabasis, K. and Katinakis, P., 2008. Characterization of spermidine and spermine synthases in *Lotus japonicus*: induction and spatial organization of polyamine biosynthesis in nitrogen fixing nodules. *Planta*, 228(1), p.37.
3. Ferraioli, S., Taté, R., Caputo, E., Lamberti, A., Riccio, A. and Patriarca, E.J., 2001. The *Rhizobium etli* *argC* gene is essential for arginine biosynthesis and nodulation of *Phaseolus vulgaris*. *Molecular plant-*

microbe interactions, 14(2), pp.250-254.

4. Omena-Garcia, R.P., Justino, G.C., Sodek, L. and de Carvalho Gonccedil, J.F., 2011. Mineral nitrogen affects nodulation and amino acid xylem transport in the Amazonian legume *Inga edulis* Mart. *International Journal of Plant Physiology and Biochemistry*, 3(12), pp.215-218.
5. Lohar, D.P., Sharopova, N., Endre, G., Penuela, S., Samac, D., Town, C., Silverstein, K.A. and VandenBosch, K.A., 2006. Transcript analysis of early nodulation events in *Medicago truncatula*. *Plant Physiology*, 140(1), pp.221-234.
6. Bianucci, E., Tordable, M.D.C., Fabra, A. and Castro, S., 2008. Importance of glutathione in the nodulation process of peanut (*Arachis hypogaea*). *Physiologia plantarum*, 134(2), pp.342-347.
7. Veličković, D., Agtuca, B.J., Stopka, S.A., Vertes, A., Koppelaar, D.W., Paša-Tolić, L., Stacey, G. and Anderton, C.R., 2018. Observed metabolic asymmetry within soybean root nodules reflects unexpected complexity in rhizobacteria-legume metabolite exchange. *The ISME journal*, 12(9), p.2335.
8. Fiantis, D., Nelson, M., Shamshuddin, J., Goh, T.B. and Van Ranst, E., 2010. Determination of the geochemical weathering indices and trace elements content of new volcanic ash deposits from Mt. Talang (West Sumatra) Indonesia. *Eurasian Soil Science*, 43(13), pp.1477-1485.

Chapter 3

Metabolic and genetic basis of bacteria-driven mineral weathering and its phylogenetic diversity at the soil-mineral interface

Dimitar Z. Epihov^{1*}, Jonathan R. Leake¹, and David J. Beerling¹

¹*Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK*

*for correspondence

Abstract

Mineral weathering in the soil environment can replenish diminishing nutrients stocks by unlocking highly insoluble mineral-bound elements. The role of soil bacteria as drivers of enhanced weathering is well-documented, but the genetic and environmental basis for the large variability existing in weathering potentials between different bacterial genera is not understood. Here, we show that several single gene mutations significantly impact the *in vitro* weathering potential of the tropical soil beta-proteobacterium *Burkholderia thailandensis*. Our functional analyses identify target weathering genes from the Entner-Doudoroff pathway, acetate metabolism and siderophore biosynthesis. Phylogenetics of the soil pools of those genes provide evidence for the role of multiple previously undescribed and elusive bacterial lineages, including a new class of Acidobacteria – *Ca. Acidipotentia* cl. nov., in weathering dynamics. Using parametrisation from our *in vitro* results we established a model to calculate the estimated weathering capacity (EWC) of single bacteria and whole communities. EWC strongly correlated with weathering *in vitro* and *in vivo* and we demonstrate its utility for biome-specific analyses. Our findings provide functional genetic and metabolic basis of enhanced weathering in soil bacteria supplying marker genes for use in soil shotgun metagenomic and metatranscriptomic analyses as well as useful targets for biotechnological modification in industrial bioleaching of mineral ores.

Weathering of silicate minerals in the soil environment increases nutrient availability, decreases N₂O emissions and increases CO₂ capture by carbonation reactions^{1,2} forming a key

part of biogeochemical cycling in the pedosphere. Soil bacteria are increasingly recognized as powerful agents of enhanced weathering^{3,4} unlocking minerals through several mechanisms. In natural systems, lithotrophic bacteria utilizing inorganic salts as source of electrons can exacerbate weathering by their effects on ecosystem nutrient cycling, specifically (1) enhanced nitrate (NO_3^-) leaching and acidification through the activity of nitrifying prokaryotic consortia^{5,6}, (2) iron (Fe) oxidation reactions dissolving Fe(II) from silicate minerals through Fe cycling bacteria^{7,8}, (3) production of corrosive sulphur (S) species through S oxidizing lineages^{9,10}. However, the majority of bacteria in soil are represented by organotrophic genera which may affect weathering by mobilizing metals from minerals through the combined action of organic acids¹¹, siderophores⁸ and Fe(III) reductive dissolution¹².

Large natural variability in bacteria-driven weathering rates exist. For example, isolation of bacterial strains from the granite forefront of the Damma Glacier in the Swiss Alps revealed very high *in vitro* solubilisation of Fe from granite by actinobacterial isolates (*Arthrobacter*, *Leifsonia*) and beta-proteobacteria (*Polaromonas*, *Janthinobacterium*)¹³. Interestingly, other Damma Glacier isolates in the same classes of actinobacteria (*Frigoribacter*) and beta-proteobacteria (*Oxalobacter*, *Paucibacter*) failed to trigger significant Fe dissolution¹³. That clearly demonstrates that class-level phylogenetic clustering does not correspond to weathering ability despite previously observed correlation between beta-proteobacteria and apatite weathering from community-profiling of soil-deposited mineral bags¹⁴. In supporting the presence of within-class variation in weathering capacity, another study found that members of the actinomycete class isolated from granite and assigned to *Kibdelosporangium* and *Actinopolyspora* exhibited the highest Fe solubilisation rates followed by intermediate rates in *Kitasatospora* with lowest rates observed in *Streptomyces* isolates¹⁵. *In vitro* analysis of >400 soil isolates from along a soil profile revealed that bacteria-driven weathering rates varied greatly from 4 to 175 μM dissolved Fe from biotite mineral¹⁶ and from 1 to 37 μM Fe solubilised from feldspar¹⁷. Variation in weathering dynamics has even been documented in different species within the same genus with rock phosphate weathering rates in solid cultures differing substantially between 3 different species from each of the alpha-proteobacterial genera *Rhizobium* and *Bradyrhizobium*¹⁸.

Currently, the genetic and metabolic basis of this variation in weathering potential in bacteria is not fully understood. Gluconic acid production has been highlighted in several studies as

important driver of acidification and chelation activities fuelling weathering^{19,20}. In two *Pseudomonas* species differing in their weathering efficiencies, high P solubilisation under P-limiting conditions were linked to increased activity of the pyrroloquinoline quinone (PQQ)-dependent glucose dehydrogenase (gdhPQQ) catalysing the oxidation of glucose to gluconic acid²¹. Similarly, transfer of PQQ synthase gene from the weathering efficient *Erwinia herbicola* to *Pseudomonas* and *Burkholderia* increased solubilisation haloes by 11-42%²². The gdhPQQ is part of the Entner-Doudoroff (ED) pathway (a metabolic alternative to the standard Embden-Meyerhof-Parnas/EMP glycolysis) in which glucose is oxidized to aldonic acids in the periplasm, subsequently transported into the cytosol and processed to 2-keto-3-deoxy-phosphogluconate (KDPG). Ultimately, KDPG is converted into pyruvate and glyceraldehyde-3-phosphate which are then converted into acetyl-CoA through standard EMP pathway finally flowing into the Krebs cycle. These findings implicate the ED pathway as key in driving weathering but how the network of genes participating in this pathway as a whole affects weathering remains unclear.

Microbially produced siderophores are known to provide bacteria with utilizable source of iron by binding to Fe(III) even at lower than yoctomolar ($<10^{-24}$ M) concentrations. However, the binding and uptake of Fe(III) already in solution is unlikely to have any pronounced effect on weathering through equilibrium dynamics as bacterial uptake is relatively small compared to the vast quantities of Fe liberated by dissolution of Fe-rich minerals. Direct siderophore-mineral interactions would stimulate weathering and such interactions have been recorded for Fe hydroxides and smectite²³ but not for primary silicate rocks such as olivine/dunite, basalt etc.

Here, we test for the direct silicate weathering effects of genes coding for (1) ED pathway and related pentose phosphate pathway enzymatic machinery ($n=9$ genes), (2) metal transporters ($n=3$ genes) and (3) siderophore production ($n=1$ gene) using a reverse genetics approach, utilizing knock-out (KO) mutants from the transposon mutant library available for the tropical soil beta-proteobacterium *Burkholderia thailandensis* E264 strain²⁴. Bacteria were incubated *in vitro* in specifically designed sterile minimal liquid medium containing 5.6 mM glucose, salts and 0.1% w/v crushed olivine with the mineral representing the only source of Fe and micronutrients.

Results

Wild type *Burkholderia thailandensis* cultures trigger a multi-elemental pattern of weathering *in vitro* stimulated by high N supply

Initial screens (first experiment) with the wild-type (WT) *Burkholderia thailandensis* E264 (BtE264) cultures using two different mineral grain sizes (60-90 μm and 90-106 μm) revealed that after a 10-day incubation period bacteria triggered multi-elemental weathering release from olivine that generated significantly larger concentrations of Fe (193-fold), cobalt (Co; 124-fold), nickel (Ni; 26-fold), manganese (Mn; 18-fold), magnesium (Mg; 15-fold), silicon (Si; 10-fold) and chromium (Cr; 3-fold) in solution than the abiotic treatment, a pattern independent on grain size ($n=6$ for each grain size and treatment; two-tailed t-test, $P<0.001$; **Figure 1a**). In the rest of our experiments, those elements were considered the 5 principal components of olivine that are liberated as a result of bacterially-mediated weathering.

We tested 3 different C:N ratios by increasing the levels of N available as ammonium sulfate supplied in the medium (C:N = 10,12,14). Weathering rates negatively correlated with medium C:N suggesting that increased N supply can stimulate weathering (**Figure 1b**). While the significantly more acidic final pH of all the BtE264 treatments at different C:N ratios (mean pH=3.91, $n=18$, SEM=0.04) may have stimulated weathering relative to their respective abiotic controls (mean pH=6.31, $n=18$, SEM=0.02), it was not responsible for the difference driven by decreasing C:N ratio as there was no significant pH difference between different C:N treatments, neither did the final pH of bacteria-inoculated samples correlate with their exerted weathering rates. However, bacterial culture density (absorbance at OD_{600nm} or simply OD₆₀₀ hereafter) revealed an opposite pattern with the greatest growth observed at the highest C:N ratio. Subsequently, we selected C:N = 14 and larger grain size for further mutant screening as to provide a more conservative view on any of the observed phenotypes.

Transporter knock-outs do not affect weathering rates and bacterial growth

In the second experiment, we utilized three transporter knock-out (KO) mutants including the magnesium transporter *corA*, a Fe:siderophore ABC transporter and the molybdenum/cobalt transporter *modA*. In each of the tests, the transporter mutants did not exhibit any significant differences from WT BtE264 in the concentration of their target elements (two-tailed t-test,

P<0.05) suggesting that membrane transport does not significantly affect weathering dynamics (**Figure 1c**). However, none of the transporter mutants exhibited any significant differences in their growth (OD₆₀₀) than WT BtE264 suggesting that alternative transporting mechanisms must exist allowing the passage of those elements into bacterial cells.

Entner-Doudoroff (ED) and pentose-phosphate (PP) metabolism significantly affects weathering potential

We used the Uniprot database and *in silico* sequence-based predictions to establish the subcellular localisation of the protein products of the target genes involved in these pathways. Membrane localisation was suggested for the flavin dinucleotide (FAD)-binding glucose dehydrogenase (gdhFAD), 2-ketogluconate reductase (trkA), gluconokinase (gntK), altronate hydrolase (uxaA) with the rest of the tested knock-out gene products assigned to the cytoplasm (**Figure 2**). In our third experiment, we tested how inactivation at different spots within the ED/PP metabolic network would affect weathering rates.

At the surface of the bacterial cell, glucose is oxidized to gluconic acid by gdhFAD. Mutation in the *gdhFAD* gene resulted in significant multi-elemental decline in weathering with the following elemental patterns: -12% in mobilised Fe and approximately -5% in Mg, Si, Mn, Co, and Ni relative to WT BtE264 (two-tailed t-test, P<0.001 for all but Si and Ni where P<0.01; **Figure 1c**). Once oxidised, the gluconic acid in the periplasm can either enter the cell or be further oxidised to 2-ketogluconic acid that also remains outside the cell with the reaction catalysed by the 2-ketogluconate reductase/gluconate 2-dehydrogenase (trkA). Mutation in the *trkA* gene resulted in a slight but significant decline of 3-4% in weathering yield of Mg, Si, Mn, Co, and Ni compared to WT BtE264 (two-tailed t-test, P<0.05; **Figure 1c**). Any non-oxidised gluconic acid can penetrate the cell through gluconate permease. Upon entry, gluconate is immediately phosphorylated by gluconokinase (gntK) to 6-phosphogluconate. Loss-of-function mutation in the *gntK* gene disabling the cell to successfully assimilate gluconic acid significantly stunted bacterial growth by 5-fold and further caused multi-elemental increase in weathering yield of ~68% in Fe and 28-33% in Mg, Si, Mn, Co and Ni in comparison to WT BtE264 (two-tailed t-test, P<0.001; **Figure 1c**).

In the cytoplasm, the 6-phosphogluconate can either enter the classic phosphorylative ED pathway and form 2-dehydro-3-deoxy-6-phosphogluconate (by the *edd* enzyme) or be re-

routed to the standard EMP glycolysis stream by conversion to 6-phosphoglucose through the 6-phosphoglucose dehydrogenase (*zwf*). KO mutation in the *zwf* gene causes a slight 3% increase in Co yield (two-tailed t-test, $P < 0.05$) and a more pronounced 22% rise in Fe dissolved from olivine relative to WT BtE264 (two-tailed t-test, $P < 0.01$; **Figure 1c**). Mutation in the gene encoding the 6-phosphofructokinase further downstream in the EMP glycolytic pathway caused only insignificant increases in weathered Co and Ni (two-tailed t-test, $P < 0.06$).

The 6-phosphogluconate not re-routed to glycolysis is converted to 2-dehydro-3-deoxy-6-phosphogluconate (KDPG; **Figure 2**). KDPG can also be formed as a result of phosphorylation of 2-dehydro-3-deoxygluconate by the *kdgK* enzyme. Interestingly, mutation in the KDPG-generating *kdgK* gene causes a 6-7% decline in weathered Mg, Si, Mn, Co, and Ni (relative to WT BtE264, two-tailed t-test, $P < 0.001$; **Figure 1c**), while loss-of-function mutation in the *eda* gene responsible for the breakdown of KDPG to glycolytic pyruvate and glyceraldehyde-3P significantly increases weathering yields in Mg, Si, and Co by 2-6% (two-tailed t-test, $P < 0.01$; **Figure 1c**). These suggest that KDPG may act as a regulatory intermediate in the ED pathway indirectly affecting weathering potential. Interestingly, we detected that the altronate hydrolase coded for by the *uxaA* gene and converting altronate (a product of glucuronate breakdown) to 2-dehydro-3-deoxygluconate also exhibited a negative effect on weathering because its KO mutants revealed greater concentrations of weathered Mn, Co, and Ni than WT cultures (two-tailed t-test, $P < 0.05$ for Mn, $P < 0.01$ for Co and Ni; **Figure 1c**).

Acetate is transported outside the cell during high C loading resulting in acetogenesis. In this process, the acetate kinase encoded in *B. thailandensis* by the *ackA* gene is key by converting acetyl-CoA generated by glycolysis to acetate which is then excreted with a counterbalancing H^+ . KO mutation in the *ackA* gene triggered a significant 3% decline in the Mg, Mn, Co, and Ni weathered from olivine relative to WT *B. thailandensis* (two-tailed t-test, $P < 0.01$).

Siderophore biosynthesis is involved in bacterially-mediated silicate weathering

The genome of *Burkholderia thailandensis* E264 only codes for the biosynthesis of the pyochelin class of siderophores also containing their related Fe-chelating salicylic acid and

dihydroaeruginosic acid. Key in this biosynthetic pathway is the *pchD* gene encoding a siderophore biosynthesis enzyme with a salicylate-AMP ligase activity²⁵. Bt *pchD* mutants revealed a multi-elemental decrease in weathered elements including 12-17% decline in Mg, Si, Mn, Co, Ni as well as larger 35% reduction in dissolved Fe in comparison to wild type bacteria (two-tailed t-test, $P < 0.001$; **Figure 1c**). Intriguingly, the *pchD* mutants also exhibited slightly yet significantly higher OD₆₀₀ readings (+7%, two-tailed t-test, $P < 0.001$) than their WT counterparts showing that those elemental reductions did not negatively impact growth. However, the mutants also generated less acidity at the end of the experiment with one-unit greater pH than in the media of their WT analogues after the 10-day incubation period.

Combined, our findings highlight the genes for siderophore biosynthesis, glucose dehydrogenases, 2-ketogluconate reductase/gluconate 2-dehydrogenase and acetate kinase (**Figure 2**) and their products as targets contributing to bacterial weathering potential.

Correcting weathering effects for OD₆₀₀

Correcting weathering effects for OD₆₀₀ of *in vitro* weathering systems involving bacterial cultures is not a common practice. None of 14 major scientific articles that have investigated the effects of different bacterial species on weathering (Supplementary Table 1) have compared and normalised weathering rates for growth measured by OD₆₀₀. One reason for not doing so is that such studies are designed to test the effect each species has on rock weathering under the same set of conditions (medium, temperature, time, shaking rate etc.) despite inherent differences in growth rates and growth potential of different lineages. Another potential pitfall of this type of normalisation, particularly relevant to our case, is that mutations often carry deleterious effects causing a decline in bacterial growth compared to its wild type baseline. Consider this example: a bacterial mutant exhibits 2-fold decreased growth due to a deleterious mutation that disrupts a process with no mechanistic links to the weathering process (e.g. translational processing) and ion concentration in solution reveal no effect on weathering relative to wild type bacteria; if corrected, however, this mutant will reveal 2-fold greater weathering rates per unit OD₆₀₀ than wild type despite its lack of effect prior to normalisation.

For instance, *gdhFAD* mutants exhibit 38.6% significantly lower OD₆₀₀ reading after 10 days of growth compared to wild type cultures (two-tailed t-test, $P < 0.001$) and prior to

normalisation they reveal 12% significant decline in weathered Fe relative to wild type (two-tailed t-test, $P < 0.001$). Consequently, because the effect on growth dwarfs that on weathering, after normalisation *gdhFAD* KO mutants reveal 23.9% greater weathering than wild type *B. thailandensis*. This assertion is mechanistically questionable because the *gdhFAD* protein participates in the catalysis of glucose to gluconic acid oxidation process, with many previous studies implicating the production of gluconic acid as a principal agent of bacterial weathering¹⁹. Similarly, *trkA* mutants exhibit 9.5% significant reduction in OD₆₀₀ (two-tailed t-test, $P < 0.001$) compared to WT BtE264 and a 3% decline in weathering-generated Fe relative to WT (two-tailed t-test, $P < 0.01$). Consequently, after normalisation *trkA* KO cultures appear to have significantly higher weathering rates than wild type by 6.2% (two-tailed t-test, $P < 0.01$). The enzyme encoded by *trkA* catalyses the production of 2-ketogluconic acid – another agent of bacterial weathering⁴⁵ thus making this finding illogical, too. The KO mutation in *gntK* causes the build-up of gluconic acid in the medium (Supplementary Note 2) which results in 68% increases in weathering-derived Fe in the medium. In terms of growth *gntK* exhibits 4.5-fold lower OD₆₀₀ readings than WT bacteria and as such the increase seems even larger as it is spread among fewer OD units.

Knock-out mutation in *gntK* significantly reduces pH relative to wild-type (down to a mean of pH 3.70; -0.70 units relative to WT; two-tailed t-test, $P < 0.001$) owing the build-up of gluconic acid (Supplementary Note 2), whereas mutation in *gdhFAD* significantly increases pH (up to a mean of pH 4.65; +0.25 units relative to WT; two-tailed t-test, $P < 0.001$) likely owing to reduced generation of gluconic acid and that of protons dissociating from it. These findings act to further contradict the weathering results reported after OD₆₀₀ normalisation in the case of *gdhFAD*. However, because the notion that the number of bacteria carries a quantitative effect on weathering cannot be ruled out, we compared the OD₆₀₀ readings of WT bacteria grown over three separate times (batches) and their weathering rates (Supplementary Note 1) and found no statistically supported links between OD₆₀₀ and weathering within the same genetic background (Pearson test, $P > 0.10$) ruling out the possibility that the number of bacteria affects weathering rates. Nevertheless, however informative or disinformative, corrections of weathering to units weathering per unit OD₆₀₀ are, for those interested, can be found integrated within the main findings presented in Figure 1 and Figure 2. Briefly, we find that after normalisation: (1) *pchD*, *gntK*, *eda* (no significant difference in OD₆₀₀ between wild-type and this background; $P > 0.10$) and *zwf* (no significant

difference in OD₆₀₀ between wild-type and this background; $P > 0.10$) show no change in their effects on weathering from before normalisation; (2) similarly, the lack of significance remains unchanged in all three transporter gene knock-out backgrounds; (3) *trkA* and *gadFDH* exhibit shifts in their effects from decline-to-increase in weathering; (4) the effects of *ackA* (no significant difference in OD₆₀₀ between wild-type and this background; $P > 0.10$) and *ddgK* (significantly lower OD₆₀₀ in this background relative to WT; $P < 0.01$) on weathering are no longer supported (Figure 1). Among those, *ackA* is the only background where growth did not significantly differ from WT and yet the relationship between *ackA* and weathering was absent after normalisation making it the only case where one could be ascertained that the deleterious effect of the mutation was not driving the apparent pattern the mutation had on weathering after normalisation. As such, readers are cautioned that any further discussions on unnormalised rates of weathering found below, particularly in the case of the *ackA* background, may be speculative.

Phylogenetics of target weathering genes from soil metagenomes reveals highly abundant elusive lineages of strong weathering potential

We utilized 3 shotgun soil metagenomes from Neotropical forests in Panama, each containing an average of 18,792,525 sequences to characterise the phylogenetics of the weathering community in soil samples by assigning taxonomy to the target genes (*gdhFAD*, *trkA*, siderophore biosynthesis genes, *ackA*) we have established in our *in vitro* weathering experiments with the tropical soil bacterium *B. thailandensis* as well as the functional analogue of *gdhFAD*-encoded FAD-dependent glucose dehydrogenase – the *gdhPQQ* gene also encoding for glucose dehydrogenase but PQQ-dependent.

Detailed search of the published literature summarised a total of 107 bacterial genera for which proof for active mineral weathering *in vitro* was available (**Supplementary Table 1**). We used those genera to estimate what percentage of soil lineages harbouring the target weathering genes were previously characterised or not (with or without weathering record, respectively).

The *gdhFAD* gene was only present in 20-30 copies representing an average of 0.00014% of each soil metagenome. Taxonomic breakdown of the gene (**Figure 3a**) revealed that over 60% of all hits were assigned to genera of known weathering ability with the beta-

proteobacterial (*Burkholderia*, *Paraburkholderia*, *Caballeronia*, *Collimonas*, *Herbaspirillum* and *Ralstonia*) and alpha-proteobacterial (*Bradyrhizobium*, *Novosphingobium*, *Mesorhizobium*, *Caulobacter*) contributing the highest proportion. Of those *gdhFAD*-harbouring lineages of no previous weathering record, members of the hard-to-culture elusive Acidobacteria phylum contributed ~11% of all hits, uncultured uncharacterised groups – 6% and delta-proteobacteria – 6%.

A total of 4400-4900 reads from each soil metagenome library were characterised as *gdhPQQ* genes amounting to a mean of 0.025% of all metagenomic hits. Nearly 90% of all *gdhPQQ*s were assigned to genera (Figure 3b) of no previous weathering record, with a major proportion, over a half of all sequences occupied by Acidobacteria (including the genera *Sulfopaludibacter* with 7%, *Solibacter* – 6%, *Koribacter* – 2%, *Terriglobus* – 1% and another 8 named genera). The most abundant group among the Acidobacteria were uncultured acidobacterial genomes assembled from metagenomic hits composing a total of 31% of all hits. Other notable bacterial groups contributing to the soil *gdhPQQ* gene pool were the bacteroidetes classes of Cytophagia (2.8%) and Flavobacteriia (0.8%), the elusive groups of Verrucomicrobia (1.3%), Gemmatimonadetes (1.3%) and alpha-proteobacteria (29% split in 18% and 11% for genera of no weathering and published weathering record, respectively), beta-proteobacteria (22% split in 3% and 19% of genera of no weathering and published weathering record, respectively), and gamma-proteobacteria (4% - all contributed by genera of known weathering activity).

Between ~3100 and 3900 metagenomic reads (a mean of 0.019% of the whole metagenome) were functionally annotated as *trkA* (2-ketogluconate reductase) or glucose 2-dehydrogenase genes – both catalysing the reversible reaction of gluconic acid to 2-ketogluconic acid at the cell surface. Approximately 56% of all hits were contributed by genera with previously published *in vitro* weathering activity (Figure 3c) – mainly consisting of alpha-proteobacterial sequences with nearest match to *Bradyrhizobium* (40.7%), *Rhizobium* (3.5%), *Mesorhizobium* (1.9%), the beta-proteobacterial *Paraburkholderia*-*Burkholderia*-*Caballeronia* close taxonomic cluster (1.2%). Among the 44% of no weathering record, significant *trkA*-contributors were Acidobacteria (15% with 14% being of uncultured acidobacteria) and newly characterised yet largely uncultured Rokubacteria (1.7%).

The *ackA* gene was present in 5900-7250 copies representing a mean of 0.036% of soil metagenomes. Around 74% of all of the *ackA* copies were assigned to genera of no weathering record with the Acidobacteria taking up 9%. The cluster of siderophore biosynthetic genes as defined by the KEGG Orthologues (KO) database amounted to 0.034% of all metagenomic reads in soil. Of these 48% belonged to genera of known weathering record including *Streptomyces* (13.6%), *Mycobacterium* (7.3%), *Bradyrhizobium* (5.8%), *Bacillus* (4.6%), *Burkholderia-Paraburkholderia-Caballeronia* clade (2.8%), *Paenibacillus* (2.5%), *Rhizobium* (1.6%) and *Pseudomonas* (1.5%). The group of no-weathering record included 1.9% Acidobacterial sequences and those of dominant (>1%) genera namely the actinobacterial *Nocardia* (1.6%), *Amycolatopsis* (1.5%), uncultured actinobacteria (1.4%), *Micromonospora* (1.3%), the alpha-proteobacterial *Rhodoplanes* (2.8%), uncultured Rhizobiales (1.3%), and *Methylobacter* (1.0%), uncultured beta-proteobacteria (1.3%) and uncultured gamma-proteobacteria (1.1%).

Next, we ranked the genus-specific weathering potential using our detailed taxonomic breakdown of metagenomic weathering genes. We found that all weathering genes *gdhFAD/gdhPQQ*, *trkA*, *ackA*, and siderophore biosynthetic genes weathering genes were present in 44 genera, of which 19 were of known weathering potential (including *Bradyrhizobium*, *Burkholderia-Paraburkholderia-Caballeronia*, *Streptomyces*, *Rhizobium*, *Pseudomonas*, *Bacillus*, *Collimonas*, *Cupriavidus*, *Frateuria*, *Dyella*-like etc.). We also found that grouping according to number of weathering genes was significantly associated with the proportion of known weathering genera confirming that the more weathering genes present in a genus, the higher the likelihood of its being already reported as a weathering agent. Nevertheless, all weathering genes were also detected in uncultured and taxonomically-undefined members of the Acidobacteria, Rhodospirillales, Rokubacteria, beta-proteobacteria, gamma-proteobacteria, Bradyrhizobiaceae, Burkholderiales. In all of those, genes encoding the synthesis of the PQQ co-enzyme were also present suggesting that their *gdhPQQ* enzymes are active.

***Candidatus* Acidipotentia, cl. nov. – a new class of the Acidobacteria phylum with pronounced acid-generating potential**

Among the abundant and uncultured species within the Acidobacteria with genomes assembled from soil metagenomes²⁶, we have identified 17 putative acidobacterial species that can convert glucose to gluconic acid, contributing a total of 14% and 8% of the soil pools of *gdhPQQ* and *gdhFAD*, respectively. Of those there were 11 species that were also able to further oxidize their periplasmatically generated gluconic acid to 2-ketogluconic acid via the *trkA*/gluconate 2-dehydrogenase gene pool to which they contributed 5% of all reads. Of the 18 gluconic acid-generating species there were 8 species that also harboured siderophore biosynthetic genes and 6 spp. were predicted to be able to produce all three – gluconic and 2-ketogluconic acids and siderophores.

Next, we used the DNA-directed RNA polymerase subunit beta' (*rpoC*) genes of six of the uncultured species and aligned them with *rpoC* sequences of known acidobacterial lineages – representing all currently described Acidobacterial classes. Using the multiple alignment of the taxonomically reliable gene marker *rpoC*²⁷, we constructed a Neighbour-Joining phylogenetic tree which showed that three of the environmental isolates (Acidobacteria bacterium 13_2_20CM_2_57_6, Acidobacteria bacterium 13_2_20CM_57_17, and Acidobacteria bacterium 13_1_40CM_2_60_7) formed a strongly supported cluster with a deep divergence node and outside all known acidobacterial classes (**Figure 3d**). Here, we named this new candidate class as *Ca. Acidipotentia* cl. nov. ('acidi' – from Latin N. m. pl. acids, 'potentia' – from Latin N. f. sing. ability, capacity; the etymology stems from the acid-producing capacity of the bacteria in this class with members containing the genes to extracellularly generate gluconic acid, 2-ketogluconic acid, and acetic acid). In this new class we also describe a new genus *Ca. Acidipotentia* gen. nov. with two species – *Ca. Acidipotentia major* ('major' – from Latin f. sing. greater) with two strains G1 and G2, and *Ca. Acidipotentia minor* ('minor' – from Latin f. sing. lesser, smaller). *Ca. A. major* G1 contains the *gdhPQQ*, *trkA*/gluconate 2-dehydrogenase and *ackA* genes contributing to each of the soil pools of those genes 1.5%, 1.1% and 0.3% of all reads, respectively. *Ca. A. major* G2 also contains the *gdhPQQ*, *trkA*/gluconate 2-dehydrogenase genes but not the *ackA* gene with 2.2% and 0.2% contributions, respectively. The genome of *Ca. A. minor* contains the *gdhPQQ* but not the *trkA*/gluconate 2-dehydrogenase and *ackA* genes, contributing 0.4% of all soil *gdhPQQ* copies. Both species contained the ability to synthesize the PQQ co-enzyme necessary for the activity of their PQQ-dependent glucose dehydrogenases with *Ca. A. major*

G1, G2 and *Ca. A. minor* containing 0.8%, 0.6% and 0.2% of all PQQ biosynthetic genes in the soil metagenomes.

Another three uncultured acidobacterial species also formed a distinct phylogenetic cluster that appeared as sister to the acidobacterial classes Solibacteres and Acidobacteriia but at low bootstrap values and more recent divergence nodes within the group. Therefore, these three species may also be a part of another previously undescribed class but their phylogenetic positioning is less clear and we did not name them, although they also contributed significantly to the gene pools of *gdhPQQ* and *trkA*/gluconate 2-dehydrogenase.

Discussion

Our experiments with *Burkholderia thailandensis* E264 and mutants in key genes show that bacteria can mediate multi-elemental incongruent solubilisation from silicates resulting from (1) strong chelating agents (siderophores; L1 class²⁸) and (2) weak chelating agents (organic acid anions; L2 class²⁸) and (3) H⁺ generation, while ruling out any significant impact on differences in bacterial growth as measured by OD₆₀₀ (Supplementary Note 1).

The genes *gdhFAD* and *trkA* appear central to the ability of Bt E264 to solubilize olivine. The gluconic acid and 2-ketogluconic acid produced through the enzymes encoded by these two genes, respectively, can promote dissolution by weak chelation (complexolysis) and by deprotonation of their carboxylic group at pH>3.86 for gluconic and pH>2.80 for 2-ketogluconic acid generating acidity (acidolysis)^{29,30}. Similarly, the negatively charged organic acid anion acetate generated by the *ackA* gene product intracellularly and exported into the medium can provide chelating activity as well as generate acidity by co-transport with counter-balancing positively charged H⁺ allowing the cell to maintain its electron potential³¹.

The significant decline in weathering yield triggered by the loss-of-function mutations in *gdhFAD* (-5% Mg) and *trkA* (-3% Mg) is much lower in its absolute value than the increase triggered by a KO mutation in the *gntK* gene (+33% Mg) causing a build-up of gluconic acid³² and probably that of 2-ketogluconic acid. The explanation behind this phenomenon may be the combined mode of complexolysis and acidolysis in gluconic acid building up extracellularly in the *gntK* background compared to the WT state in which most gluconate is acquired and metabolized by the cell with little remaining as gluconate and converted into 2-

ketogluconic acid outside the cell thus creating weathering conditions predominated by acidolysis.

Interestingly, the phenotype of *gntK* mutants is consistent with that of growth arrest typically detected in *gntK* mutant *E. coli* supplied only with gluconate as its sole C source³². That suggests that most of the glucose supplied in our minimal medium is converted into gluconate and acquired and metabolized through the semi-phosphorylative ED pathway³³ rather than being acquired through the phosphotransferase system and direct phosphorylation of glucose to 6-phosphoglucose in the classic ED pathway. This defines that the direction of reversible reactions catalysed by enzymes such as *trkA* will be mainly from gluconate to 2-ketogluconic acid just like that of *zwf* will be mainly from 6-phosphogluconate to 6-phosphoglucose. Interruption in the latter in the *zwf* mutant background will be expected to increase the availability of 6-phosphogluconate by preventing its re-routing to EMP glycolysis. 6-phosphogluconate is a key intermediate in the erythrose-generating pentose phosphate pathway. Erythrose is used as a precursor of aromatic amino acids which are in their own right precursors for chorismate in siderophore biosynthesis. Indeed, we have recorded that *zwf* mutants exhibit a modest 3% increase in Co but a much larger 22% increase in Fe consistent with rise in the production of siderophores.

Our weathering experiments revealed that accumulation of 2-dehydro-3-deoxy-6-phosphogluconate (KDPG) in the *eda* mutant background and depletion in the *kdgA* background had opposite effects on weathering suggesting some regulatory function of KDPG within the ED pathway. Indeed, previous work indicates that KDPG positively regulates the ED pathway by binding to ED pathway repressor *hexR* and dissociating it from DNA allowing up-regulated expression of *edd*, *glc* and *zwf*³⁴ as well as by binding and inactivating the pyruvate metabolism repressor *RccR*³⁵. While inactivation of the metabolic repressors *hexR* and *RccR* can stimulate overall bacterial metabolism and thus generate more CO₂, acetate and Krebs cycle organic acids, it may limit the amount of 6-phosphogluconate by re-routing it to generation of 6-phosphoglucose metabolized in the EMP glycolytic pathway by the means of increasing the expression of *zwf* (**Figure 2**). Indeed, in our *eda* mutants accumulating higher levels of KDPG acting to alleviate the repression over metabolism – the Mg, Si and Co weathered from olivine increased 2-6% (two-tailed t-test, *P*<0.01) but Fe yields decreased by 25% (two-tailed t-test, *P*=0.06).

The decrease in olivine weathering caused by knock-out of the pyochelin biosynthetic gene *pchD* is substantial. As previously highlighted siderophore biosynthesis is dependent upon the PPP intermediate erythrose as a precursor for aromatic amino acids. Amino acid biosynthesis also requires adequate N supply. Therefore, it is plausible that the weathering-promoting effects of lower C:N ratio that we recorded are due to increasing the N availability to siderophore biosynthesis. In addition, greater N supply may promote the production of Krebs cycle organic acids by positive regulation of the Krebs cycle³⁶. Experiments with the gluconic acid-producing fungus *Aureobasidium pullulans* also demonstrated that higher N supply may promote the amount of generated gluconic acid³⁷.

Our analyses (**Figure 1c, Figure 2**) implicate *pchD* (KO effect: -35% Fe, -13% Mg) or other siderophore synthesis genes, *gdhFAD* (-12% Fe, -5% Mg) or its functional analogue *gdhPQQ*, *trkA* (-3% Mg), and *ackA* (-3% Mg) as target genes **with direct effects on weathering that can be deployed** for estimating the weathering ability of single bacteria as well as that of whole microbial communities. To validate the predicting power of these genes we utilized a quantitative study¹⁸ on the ability of different legume-nodulating bacteria to dissolve insoluble phosphate mineral sources. This study was chosen because the bacteria used in it are obtained from a taxonomically curated collection and identified down to the strain level with 5 out of the 8 bacterial strains studied having fully sequenced publicly available genomes. Using the number of target genes (siderophore genes are counted as binary data – 0/1 for absence/presence) multiplied by the Mg weathering coefficients attained from our experimental data (0.520 for siderophores, 0.208 for *gdh*, 0.125 for *trkA*/gluconate 2-dehydrogenase and 0.125 for *ackA*) we obtained estimated weathering capacity (EWC) values. The calculated EWC for the 5 bacteria (**Figure 4a**) significantly correlated with: (1) the recorded CaHPO₄ solubilisation haloes in solid media (Pearson test, $P < 0.05$, $r = 0.887$), (2) the solubilised P weathered from CaHPO₄ in liquid cultures (Pearson test, $P < 0.05$, $r = 0.914$) and with (3) the P dissolved from strengite minerals in liquid cultures – strongly insoluble iron phosphate minerals (FePO₄·2H₂O, Pearson test, $P < 0.01$, $r = 0.968$). EWC of whole microbial communities using coupled soil/mineral interfaces from our field experiments in Panama revealed a significant strong **correlation** with **in vivo** weathering rates (**Figure 4b**). Using publicly available metagenomes, we estimated the EWC of the soil communities of 4 different biomes showing a descending pattern in weathering ability (**Figure 4c**) mirroring

latitudinal gradients in precipitation, temperature and net productivity pumping organics to the communities belowground.

Highly weathered oxisolic soils in tropical forests contain little primary minerals (<1%) but most of their phosphate is present in highly insoluble mineral P forms, mainly as Al-P and Fe-P secondary minerals (such as strengite, variscite and their amorphous analogues)^{38,39}. Because net primary productivity of tropical forests is limited by P availability^{39,40}, the study of the EWC of the soil microbial community in such biomes is of high relevance. Our phylogenetic analysis of the target weathering genes in 3 tropical forest soil metagenomes demonstrated large phylogenetic diversity in weathering potent groups. Weathering potential in soil appears to be a function of the combined action of hundreds of bacterial species distributed among various bacterial phyla and classes including many poorly studied, uncultured and elusive lineages among the Acidobacteria, Rokubacteria, Glassbacteria, Verrucomicrobia etc. Although a certain level of functional redundancy exists, specialisation was also evident. For instance, some groups exhibited high EWC potential based on containing and contributing a disproportionate number of genes to the soil pools of *gdhPQQ* relative to that of siderophores (gluconic acid-based weathering agents) including acidobacterial candidate class of Acidipotentia (62:1), other acidobacteria such as *Ca. Solibacter* (21:1), *Terriglobus* (273:1), *Ca. Sulfopaludibacter* (26:1), the alpha-proteobacterial *Phenylobacterium* (81:1), *Sphingobium* (85:1), and *Rhodopila* (8:1). In contrast, other lineages had EWC mainly based on contributing a disproportionate number of copies to the soil pool of siderophore biosynthetic genes relative to that of *gdhPQQ* (siderophore-based weathering agents) such as the actinobacterial *Streptomyces* (317:1), *Amycolatopsis* (54:1), *Actinomadura* (*gdhPQQ* absent), *Arthrobacter* (*gdhPQQ* absent), the firmicute genus of *Bacillus* (125:1), uncultured Rhizobiales (93:1) etc. From the metagenomic analyses also becomes apparent that some dominant *gdhPQQ*-contributing lineages can synthesize their own internal supply of PQQ co-enzyme to activate their *gdhPQQ* apoenzyme (*Ca. Acidipotentia*, *Terriglobus*, *Rhodopila*), while others can only rely on external supplies of PQQ co-enzyme as they lack the necessary PQQ synthesis genes (*Ca. Sulfopaludibacter*, *Ca. Solibacter*, *Phenylobacterium*). Ecologically, these data present an interesting insight into the ability of moderate acidophiles such as *Terriglobus*⁴¹, *Rhodopila*⁴² and Acidipotentia candidate spp. (also likely acidophilic) to lower pH favouring their acidophilic lifestyle by producing gluconic acids dependent on sugar supply and of other acidophiles such as *Ca.*

Solibacter⁴³ and *Ca. Sulfopaludibacter*⁴⁴ by producing gluconic acids dependent on sugar supply and external PQQ co-enzymes thus enabling such moderately acidophilic lineages to dominate over more neutrophilic bacteria. Study on the wheat rhizosphere have concluded that 2-ketogluconic acid amounted for up to 22% all rhizosphere products⁴⁵ suggesting that sugar limitation is unlikely and making N availability (as we have demonstrated) and PQQ synthesis and supply highly probable limiting factors.

In conclusion, we have quantitatively characterised the effects of genetic knock-outs in a soil native bacterium in relation to the process of mineral weathering. Our analyses allow informed predictions of estimated weathering potential of single bacteria and whole microbial communities by using a combination of target genes and weathering coefficients. Consequently, this could enable the integration of the vast publicly available metagenomic datasets into biome-specific biogeochemical model estimates of weathering rates based on labile C supply. Furthermore, the large phylogenetic diversity existing in weathering potent microorganisms in soil can be used for targeted improvement of nutrient availability in agroecosystems by exploring the enormous resource of poorly characterised bacterial lineages and engineering improved PQQ supply.

Methods and materials

Handling of bacteria and generating inocula

The tropical soil beta-proteobacterium *Burkholderia thailandensis* strain E264 – wild type (WT) and its mutants were purchased from the transposon mutant collection created by the Manoil Lab at University of Washington, US. Upon arrival bacteria were streaked on LB plates and inoculated in LB broth + 10% glycerol for long-term storage at -80°C. Overnight cultures in 15 ml LB were brought to absorbance reading (OD₆₀₀) of 0.680 and diluted or concentrated as necessary. Resulting 1 ml initial inocula were centrifuged at 5000 rpm for 5 min. Subsequently, the supernatant was removed and the pellet was washed with 1ml nuclease-free sterile H₂O and re-suspended. The procedure was repeated 2 times to remove any remaining LB broth. A 10 µl aliquot of the resulting 1ml final fully re-suspended inoculum was used for inoculation in the *in vitro* weathering experiments.

In vitro weathering experiments

We created a specifically-designed minimal medium (Dunite grain rock abbreviated DGrock) containing 1 g/L D-glucose, 0.03 g/L K_2HPO_4 , 0.05 g/L NaCl, 0.002 g/L $MgSO_4 \cdot 7H_2O$ and 0.1347 g/L $(NH_4)_2SO_4$ dissolved in double-distilled water (this is the recipe for C:N=14; for C:N=12 and 10 the $(NH_4)_2SO_4$ amount was amended to 0.1572 and 0.1885 g/L, respectively). The rationale behind the medium was to contain only miniscule amounts of Mg as to allow the initiation of growth before weathering of the Mg-rich olivine commenced (that added amount of Mg was deducted from the final Mg weathering rates). Also, the medium contained no soluble forms of Fe and micronutrients with those only available through dissolution of the olivine mineral. Olivine grains were through hammer-crushing of dunite silicate rock (92% olivine mineral) from Åheim, Norway. The resulting grain mixture was sieved to desired sizes (70-90 and 96-106 μm grain diameter) and washed in 0.1N ultra-pure HCl (Normatom®, VWR chemicals) to remove any exchangeable ions⁴⁶ and dust attached to the surface of the grains and subsequently washed up to 10-15 times with double-distilled H_2O to remove any remaining acid and until the solution was fully transparent with no cloudiness associated with dust and attached ions. The resulting grains were then fully dried for several hours in a 60°C oven, mixed thoroughly and used for experiments. To equalise the amounts of olivine added to each flask (replicate), 0.07 g of olivine was added to a 250 ml Erlenmeyer flask, autoclaved and then 70 ml of autoclaved medium was poured in (to a final amount of 0.1% (w/w) olivine) using an autoclaved volumetric flask. The resulting 70 ml-filled flasks containing olivine were inoculated with 10 μl of the final bacterial inoculum and plugged with pre-autoclaved lid composed of aluminium foil and non-absorbing cotton wool. Each of the experimental flasks (replicates) were randomly placed on a horizontal shaker and incubated for 10 days, at temperature of 20°C and at 130 rpm. Abiotic controls were prepared as described but instead of 10 μl inoculum, they were mock-inoculated with 10 μl double-distilled H_2O . Growth assessment at the end of the experiment was spectrophotometrically assessed at OD₆₀₀ using 3 ml of the final culture.

Culture filtrates for ICP-MS

An aliquot of 13 ml from each final 10-day culture were obtained in sterile 15 ml tubes, centrifuged at 4200 rpm for 5 min to pellet the bacterial cells. The supernatant was transferred to a fresh 15 ml tube by filtering with a 20 ml syringe fitted with a sterile 0.22 μm Millex®-

syringe filter. 9.41 ml of the resulting cell-free filtrate was mixed with 0.59 ml 34% ultra-pure HCl to a final 2% HCl concentration (to prevent abiotic precipitation of Fe and Si⁴⁶) and sent for ICP-MS analysis. The latter was carried out at the University of Nottingham, UK using Rh, Ge, Sc as internal standards to correct for any drift in light bivalent metals of low atomic weight. The remaining amount of cell-free filtrate was used for pH measurement.

Soil metagenomes and metagenomic analysis

Soil DNA was extracted from 0.25 g homogenized soil using MoBio PowerSoil DNA extraction kit and instruction therein. Soil was obtained from topsoil samples (0-10cm) collected in the rooting zones of young legume trees in secondary tropical forests as part of the Agua Salud Project managed by the Smithsonian Tropical Research Institute in Colón Province, Panama (latitude: 9.187, longitude: -79.79). Extracted soil DNA was prepared using the Nextera XT kit and protocol, and sequenced on 1 lane of the Illumina HiSeq 4500 150bp paired-ends (together with another 12 mineral-associated DNA libraries) at the Edinburgh Genomics Institute, UK. Next, mate pairs FASTA files were merged and submitted for automated shotgun metagenomic analysis in the MG-RAST server.

Gene searching and taxonomic breakdown

All reads from each of the identified target weathering genes were downloaded as FASTA files directly from the Subsystem database-annotated metagenomes using the MG-RAST website interface. The downloaded FASTA files were then submitted for further taxonomic breakdown using the Blast2GO 5.0 Pro and their Cloud-based BLAST server. All Blast2GO hits were defined at the species level without any higher taxonomic ranking. That was obtained manually using the Taxonomy Summary facility available at <https://www.genome.jp/tools-bin/taxsummary>.

Sub-cellular localisation

The sub-cellular localisation of gene products (as illustrated in Figure 2) were either found in the Uniprot database or predicted using the Cell-PLoc 2.1 software, and particularly its sub-

domain Gneg-mPLOC designed specifically for in silico predictions in Gram negative bacteria such as *B. thailandensis*.

Phylogenetic tree construction

Multiple alignment of acidobacterial DNA-directed RNA polymerase beta' subunit (*rpoC*) gene sequences was performed using MUSCLE in the MEGA7 software environment. The multiple alignments were then utilized to construct a phylogenetic tree with the following parameters: Neighbour-Joining method and Kimura-2-parameter distance with bootstrap $n=5000$ replicates. Taxonomic binning through the *rpoC* gene as a barcode has been previously successfully applied with the same parameters in the classification of *Leuconostoc* spp. and other related genera²⁷ agreeing well with other popular markers such as 16S rRNA, *gyrB*, and *dnaA*²⁷.

Download of publicly available genome assemblies and EWC modelling

Genomes required for *rpoC* gene phylogenetic analysis or for analysis of the estimated weathering capacity (EWC) were downloaded from the NCBI Genome Assembly database, whereas metagenomes were obtained through the MG-RAST database. EWC modelling in single bacteria was achieved by multiplying the number of weathering genes by their relative proportion of contribution to Mg weathering from olivine as shown by our *in vitro* experiments with their knock-outs relative to WT bacteria (weathering coefficient ratio was 5:3:15:3 for the effects of *gdhFAD*/*gdhPQQ* : *trkA*/gluconate 2-dehydrogenase : siderophore genes : *ackA*; for single bacteria siderophore production was counted as binary data – presence/absence; the effect of *gdhFAD* we report here is assumed to be the same as that of *gdhPQQ*; the same goes for *trkA* and gluconate 2-dehydrogenase). For modelling the EWC of soil metagenomes, a similar method was used in which the proportion of each of those gene classes (normalized by *rpoC* copies in each metagenome because each bacterium contains only one *rpoC* thus surrogate for normalizing for the actual number of bacterial genomes in each metagenome) was multiplied by the aforementioned weathering coefficients. For establishing the metagenomic abundance of siderophores we used the Level 2 Subsystem database entry under the name of “Siderophores” available for each metagenome in MG-RAST, all other target genes were defined as function (Level 4) entries for each metagenome in the MG-RAST server.

References

- Beerling, D.J., Leake, J.R., Long, S.P., Scholes, J.D., Ton, J., Nelson, P.N., Bird, M., Kantzas, E., Taylor, L.L., Sarkar, B. and Kelland, M., 2018. Farming with crops and rocks to address global climate, food and soil security. *Nature plants*, 4(3), p.138.
- Kantola, I.B., Masters, M.D., Beerling, D.J., Long, S.P. and DeLucia, E.H., 2017. Potential of global croplands and bioenergy crops for climate change mitigation through deployment for enhanced weathering. *Biology letters*, 13(4), p.20160714.
- Chen, W., Wang, Q., He, L. and Sheng, X., 2016. Changes in the weathering activity and populations of culturable rock-weathering bacteria from the altered purple siltstone and the adjacent soil. *Geomicrobiology journal*, 33(8), pp.724-733.
- Uroz, S., Calvaruso, C., Turpault, M.P. and Frey-Klett, P., 2009. Mineral weathering by bacteria: ecology, actors and mechanisms. *Trends in microbiology*, 17(8), pp.378-387.
- Szwerinski, H., Arvin, E. and Harremoës, P., 1986. pH-decrease in nitrifying biofilms. *Water Research*, 20(8), pp.971-976.
- Epihov, D.Z., Batterman, S.A., Hedin, L.O., Leake, J.R., Smith, L.M. and Beerling, D.J., 2017. N₂-fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?. *Proceedings of the Royal Society B: Biological Sciences*, 284(1860), p.20170370.
- Leslie, K., Sturm, A., Stotler, R., Oates, C.J., Kyser, T.K. and Fowle, D.A., 2015. *Marinobacter* bacteria associated with a massive sulphide ore deposit affect metal mobility in the deep subsurface. *Geochemistry: Exploration, Environment, Analysis*, 15(4), pp.319-326.
- Pollmann, K., Kutschke, S., Matys, S., Kostudis, S., Hopfe, S. and Raff, J., 2016. Novel biotechnological approaches for the recovery of metals from primary and secondary resources. *Minerals*, 6(2), p.54.
- Johnson, D.B., Kanao, T. and Hedrich, S., 2012. Redox transformations of iron at extremely low pH: fundamental and applied aspects. *Frontiers in microbiology*, 3, p.96.
- Hamilton, W.A., 1985. Sulphate-reducing bacteria and anaerobic corrosion. *Annual review of microbiology*, 39(1), pp.195-217.
- Varadachari, C., Barman, A.K. and Ghosh, K., 1994. Weathering of silicate minerals by organic acids II. Nature of residual products. *Geoderma*, 61(3-4), pp.251-268.
- Bennett, P.C., Rogers, J.R., Choi, W.J. and Hiebert, F.K., 2001. Silicates, silicate weathering, and microbial ecology. *Geomicrobiology Journal*, 18(1), pp.3-19.
- Frey, B., Rieder, S.R., Brunner, I., Plötze, M., Koetzsch, S., Lapanje, A., Brandl, H. and Furrer, G., 2010. Weathering-associated bacteria from the Damma glacier forefield: physiological capabilities and impact on granite dissolution. *Appl. Environ. Microbiol.*, 76(14), pp.4788-4796.
- Lepleux, C., Turpault, M.P., Oger, P., Frey-Klett, P. and Uroz, S., 2012. Correlation of the abundance of betaproteobacteria on mineral surfaces with mineral weathering in forest soils. *Appl. Environ. Microbiol.*, 78(19), pp.7114-7119.
- Abdulla, H., 2009. Bioweathering and biotransformation of granitic rock minerals by actinomycetes. *Microbial ecology*, 58(4), pp.753-761.
- Wang, Q., Wang, R.R., He, L.Y., Lu, J.J., Huang, Z. and Sheng, X.F., 2014. Changes in weathering effectiveness and community of culturable mineral-weathering bacteria along a soil profile. *Biology and fertility of soils*, 50(7), pp.1025-1034.
- Wang, Q., Cheng, C., He, L., Huang, Z. and Sheng, X., 2014. Characterization of depth-related changes in bacterial communities involved in mineral weathering along a mineral-rich soil profile. *Geomicrobiology Journal*, 31(5), pp.431-444.
- Marra, L.M., Oliveira, S.M.D., Soares, C.R.F.S. and Moreira, F.M.D.S., 2011. Solubilisation of inorganic phosphates by inoculant strains from tropical legumes. *Scientia Agricola*, 68(5), pp.603-609.
- Lin, T.F., Huang, H.I., Shen, F.T. and Young, C.C., 2006. The protons of gluconic acid are the major factor responsible for the dissolution of tricalcium phosphate by *Burkholderia cepacia* CC-A174. *Bioresource Technology*, 97(7), pp.957-960.
- Chen, W., Luo, L., He, L.Y., Wang, Q. and Sheng, X.F., 2016. Distinct mineral weathering behaviors of the novel mineral-weathering strains *Rhizobium yantingense* H66 and *Rhizobium etli* CFN42. *Appl. Environ. Microbiol.*, 82(14), pp.4090-4099.
- Buch, A., Archana, G. and Kumar, G.N., 2008. Metabolic channeling of glucose towards gluconate in phosphate-solubilizing *Pseudomonas aeruginosa* P4 under phosphorus deficiency. *Research in microbiology*, 159(9-10), pp.635-642.
- Liu, S.T., Lee, L.Y., Tai, C.Y., Hung, C.H., Chang, Y.S., Wolfram, J.H., Rogers, R. and Goldstein,

- A.H., 1992. Cloning of an *Erwinia herbicola* gene necessary for gluconic acid production and enhanced mineral phosphate solubilization in *Escherichia coli* HB101: nucleotide sequence and probable involvement in biosynthesis of the coenzyme pyrroloquinoline quinone. *Journal of Bacteriology*, 174(18), pp.5814-5819.
23. Ferret, C., Sterckeman, T., Cornu, J.Y., Gangloff, S., Schalk, I.J. and Geoffroy, V.A., 2014. Siderophore-promoted dissolution of smectite by fluorescent *Pseudomonas*. *Environmental microbiology reports*, 6(5), pp.459-467.
24. Gallagher, L.A., Ramage, E., Patrapuvich, R., Weiss, E., Brittnacher, M. and Manoil, C., 2013. Sequence-defined transposon mutant library of *Burkholderia thailandensis*. *MBio*, 4(6), pp.e00604-13.
25. Serino, L., Reimann, C., Visca, P., Beyeler, M., Chiesa, V.D. and Haas, D., 1997. Biosynthesis of pyochelin and dihydroaeruginosic acid requires the iron-regulated *pchDCBA* operon in *Pseudomonas aeruginosa*. *Journal of bacteriology*, 179(1), pp.248-257.
26. Butterfield, C.N., Li, Z., Andeer, P.F., Spaulding, S., Thomas, B.C., Singh, A., Hettich, R.L., Suttle, K.B., Probst, A.J., Tringe, S.G. and Northen, T., 2016. Proteogenomic analyses indicate bacterial methylotrophy and archaeal heterotrophy are prevalent below the grass root zone. *PeerJ*, 4, p.e2687.
27. Chelo, I.M., Ze-Ze, L. and Tenreiro, R., 2007. Congruence of evolutionary relationships inside the *Leuconostoc–Oenococcus–Weissella* clade assessed by phylogenetic analysis of the 16S rRNA gene, *dnaA*, *gyrB*, *rpoC* and *dnaK*. *International journal of systematic and evolutionary microbiology*, 57(2), pp.276-286.
28. Hassler, C.S., Schoemann, V., Nichols, C.M., Butler, E.C. and Boyd, P.W., 2011. Saccharides enhance iron bioavailability to Southern Ocean phytoplankton. *Proceedings of the National Academy of Sciences*, 108(3), pp.1076-1081.
29. Duff, R.B., Webley, D.M. and Scott, R.O., 1963. Solubilization of minerals and related materials by 2-ketogluconic acid-producing bacteria. *Soil Science*, 95(2), pp.105-114.
30. Van Rompaey, K., Van Ranst, E., Verdoodt, A. and De Coninck, F., 2007. Use of the test-mineral technique to distinguish simple acidolysis from acido-complexolysis in a Podzol profile. *Geoderma*, 137(3-4), pp.293-299.
31. Wolfe, A.J., 2005. The acetate switch. *Microbiol. Mol. Biol. Rev.*, 69(1), pp.12-50.
32. Robin, A. and Kepes, A., 1975. Inducible gluconate permease in a gluconate kinase-deficient mutant of *Escherichia coli*. *Biochimica et Biophysica Acta (BBA)-Biomembranes*, 406(1), pp.50-59.
33. Ahmed, H., Ettema, T.J., Tjaden, B., Geerling, A.C., van der Oost, J. and Siebers, B., 2005. The semi-phosphorylative Entner–Doudoroff pathway in hyperthermophilic archaea: a re-evaluation. *Biochemical Journal*, 390(2), pp.529-540.
34. Daddaoua, A., Krell, T. & Ramos, J. L., 2009. Regulation of glucose metabolism in *Pseudomonas*. The phosphorylative branch and Entner–Doudoroff enzymes are regulated by a repressor containing a sugar isomerase domain. *Journal of Biological Chemistry*, 284(32), pp.21360-21368.
35. Campilongo, R., Fung, R.K., Little, R.H., Grenga, L., Trampari, E., Pepe, S., Chandra, G., Stevenson, C.E., Roncarati, D. and Malone, J.G., 2017. One ligand, two regulators and three binding sites: How KDPG controls primary carbon metabolism in *Pseudomonas*. *PLoS genetics*, 13(6), p.e1006839.
36. Shimizu, K., 2014. Regulation systems of bacteria such as *Escherichia coli* in response to nutrient limitation and environmental stresses. *Metabolites*, 4(1), pp.1-35.
37. Anastassiadis, S., Aivasidis, A. and Wandrey, C., 2003. Continuous gluconic acid production by isolated yeast-like mould strains of *Aureobasidium pullulans*. *Applied Microbiology and Biotechnology*, 61(2), pp.110-117.
38. Chacon, N., Flores, S. and Gonzalez, A., 2006. Implications of iron solubilization on soil phosphorus release in seasonally flooded forests of the lower Orinoco River, Venezuela. *Soil Biology and Biochemistry*, 38(6), pp.1494-1499.
39. Townsend, A.R., Cleveland, C.C., Asner, G.P. and Bustamante, M.M., 2007. Controls over foliar N: P ratios in tropical rain forests. *Ecology*, 88(1), pp.107-118.
40. Batterman, S.A., Hedin, L.O., Van Breugel, M., Ransijn, J., Craven, D.J. and Hall, J.S., 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature*, 502(7470), p.224.
41. Eichorst, S.A., Breznak, J.A. and Schmidt, T.M., 2007. Isolation and characterization of soil bacteria that define *Terriglobus* gen. nov., in the phylum Acidobacteria. *Appl. Environ. Microbiol.*, 73(8), pp.2708-2717.
42. Imhoff, J.F., Rahn, T., Künzel, S. and Neulinger, S.C., 2018. New insights into the metabolic potential of the phototrophic purple bacterium *Rhodospira globiformis* DSM 161 T from its draft genome sequence and evidence for a vanadium-dependent nitrogenase. *Archives of microbiology*, 200(6), pp.847-857.
43. Ward, N.L., Challacombe, J.F., Janssen, P.H., Henrissat, B., Coutinho, P.M., Wu, M., Xie, G., Haft, D.H., Sait, M., Badger, J. and Barabote, R.D., 2009. Three genomes from the phylum Acidobacteria provide insight into the lifestyles of these microorganisms in soils. *Appl. Environ. Microbiol.*, 75(7), pp.2046-2056.
44. Hausmann, B., Pelikan, C., Herbold, C.W., Köstlbacher, S., Albertsen, M., Eichorst, S.A., Del Rio,

- T.G., Huemer, M., Nielsen, P.H., Rattei, T. and Stingl, U., 2018. Peatland Acidobacteria with a dissimilatory sulfur metabolism. *The ISME journal*, 12(7), p.1729.
45. Moghimi, A., 1977. *Significance of 2-ketogluconic acid in dissolution of phosphate minerals by rhizosphere products*/by Azar Moghimi (Doctoral dissertation).
46. Yu, Y., Sheng, X., He, L. and Huang, Z., 2016. Linkage between culturable mineral-weathering bacteria and their weathering effectiveness along a soil profile. *Geomicrobiology Journal*, 33(1), pp.10-19.

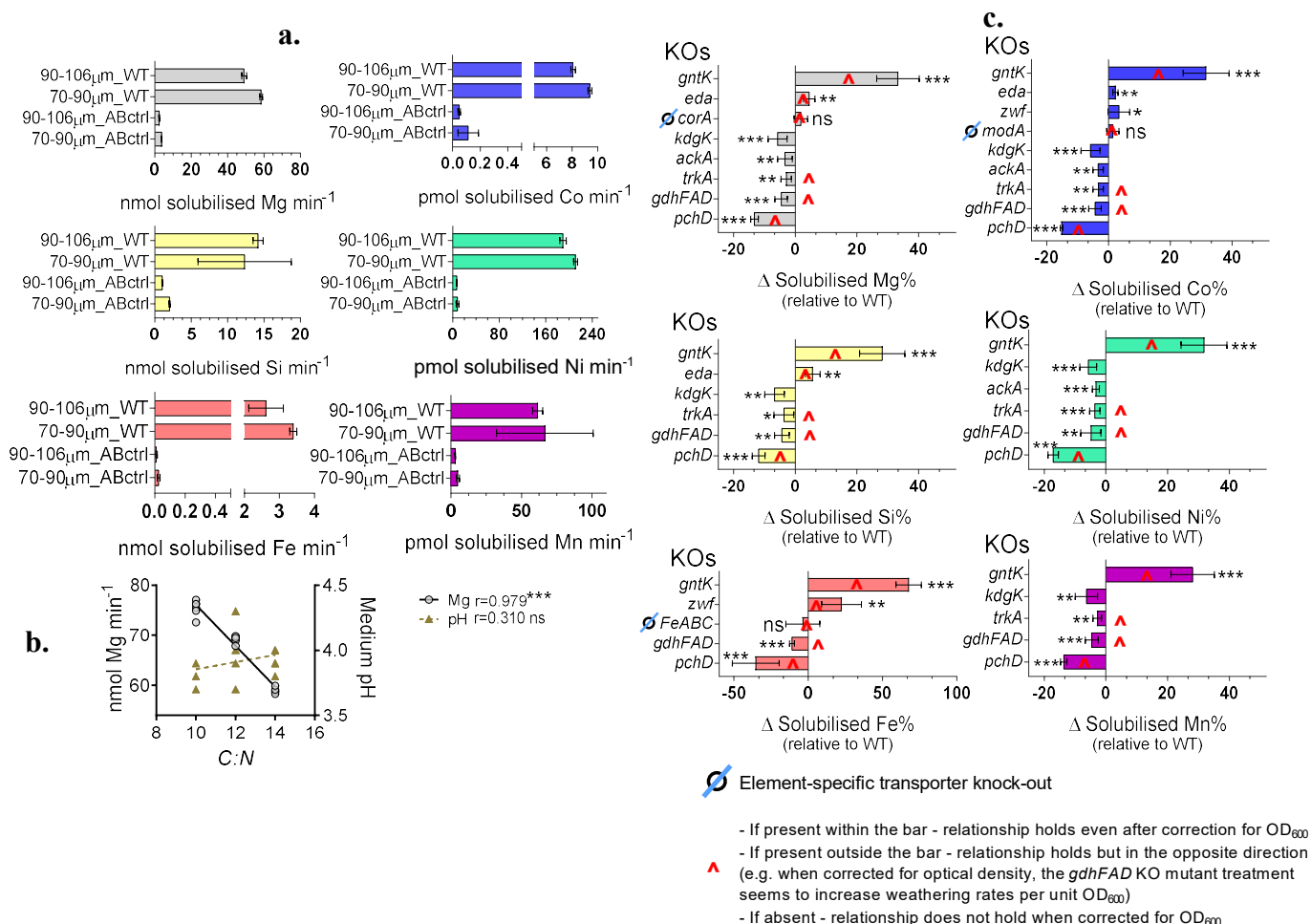


Figure 1. Reverse genetics analyses of the multi-elemental weathering potential of the soil bacterium *Burkholderia thailandensis* E264 highlights siderophores and Entner-Doudoroff pathway target genes. **a.** Establishing the *in vitro* BtE264 weathering potential relative to abiotic controls using different grain sizes; **b.** Weathering potential correlates with decreasing medium C:N ratio (increasing N supplied as ammonium sulfate) but not with pH in WT bacterial cultures; **c.** Positive effectors of weathering appear to be encoded by *pchD*, *gdhFAD*, *trkA*, *ackA*, *kdgK* while negative effectors include *gntK*, *eda*, *zwf*. The *r* values in **b.** are for Pearson correlation test; statistical tests in **c.** (** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$) are for two-tailed t-test comparing the mutant with its appropriate batch-specific wild type (WT) treatment. Abbreviations: ABctrl – abiotic control; KOs – genetic knock-outs from the transposon mutant library available for BtE264; *gntK* – gluconate kinase; *eda* – 2-dehydro-3-deoxy-6-phosphogluconate aldolase; *zwf* – 6-phosphogluconate dehydrogenase; *corA* – magnesium transporter; *modA* – molybdenum/cobalt transporter; *FeABC* – Fe compound ATP-binding cassette (ABC) transporter; *kdgK* – 2-dehydro-3-deoxygluconate kinase; *trkA* – 2-ketogluconate reductase/gluconate 2-dehydrogenase; *gdhFAD* – FAD-dependent glucose dehydrogenase; *pchD* – pyochelin siderophore biosynthetic protein D. $n=6$ bacterial cultures in each experiment.

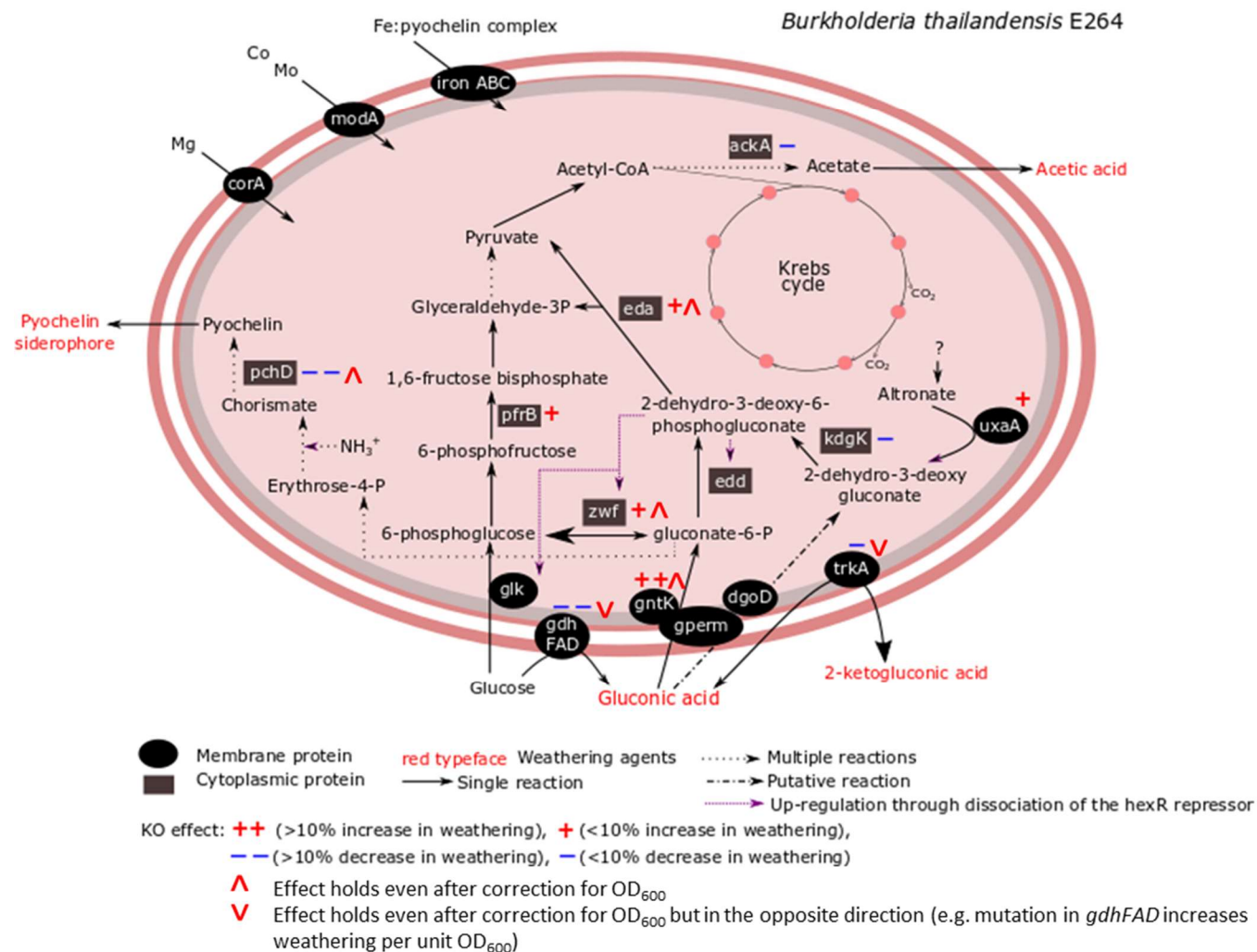


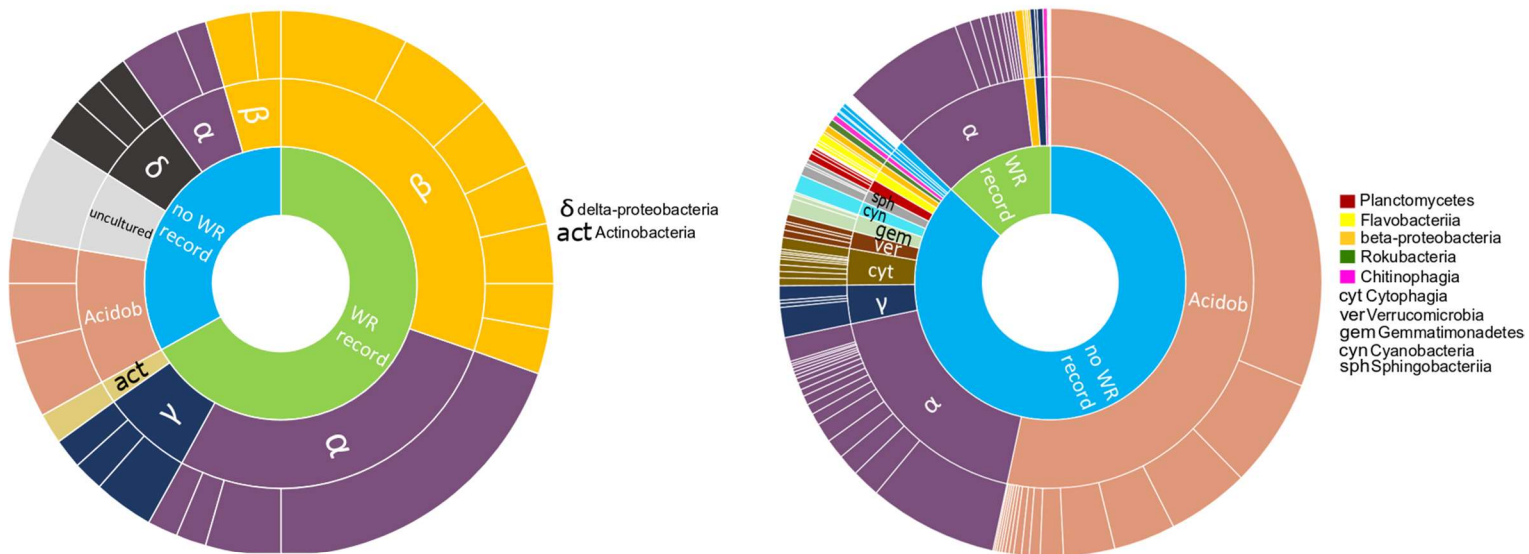
Figure 2. Pathway network effects in *Burkholderia thailandensis* reveals direct (*gdhFAD*, *trkA*, *ackA*, *pchD*, *gntK*) and indirect (*kdgK*, *eda*, *zwf*) mechanisms regulating weathering with the former involved in generating acidifying or chelating activity outside the cell and the latter modulating the response of the former from inside the cell.

a.

FAD-binding glucose dehydrogenase (gdhFAD)

b.

PQQ-binding glucose dehydrogenase (gdhPQQ)



c.

Soil metagenomic pool of
2-ketogluconate reductase (*trkA*)

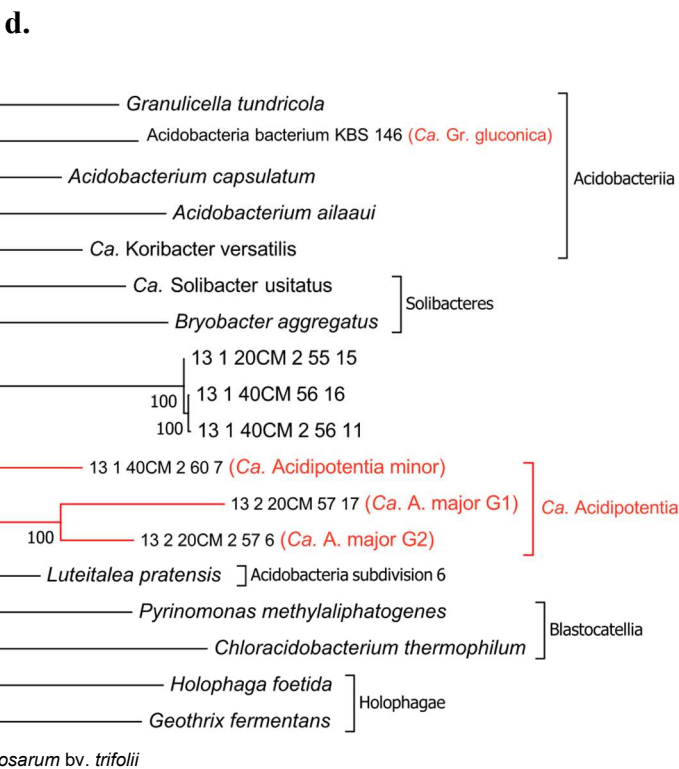
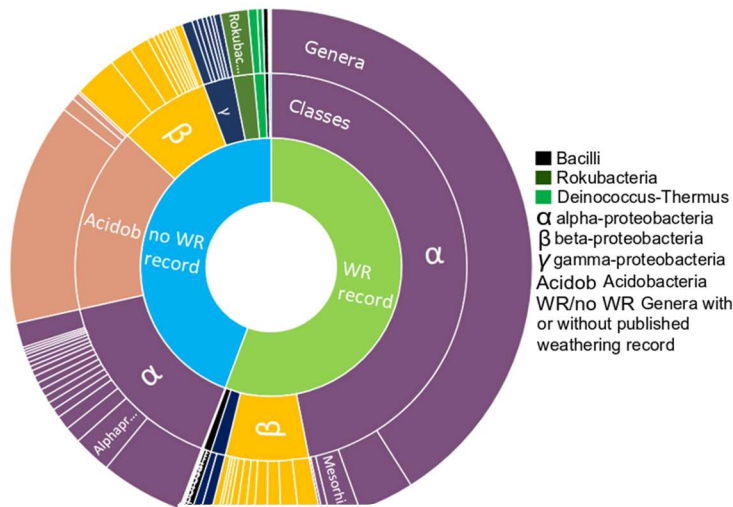


Figure 3. Phylogenetics of the target weathering genes (a – *gdhFAD*, b – *gdhPQQ*, c – *trkA*) in tropical soil communities indicates multiple weathering potent lineages of unstudied, uncultured and/or elusive bacteria including *Ca. Acidipotentia* – cl. nov. and *Ca. Granulicella gluconica* sp. nov. within the Acidobacterial phylum with their exact phylogenetic positioning (d). ‘Weathering record’ or ‘no weathering record’ are based on a list of genera with recorded weathering activity *in vitro* (summarised from published literature in Supplementary Table 1). The phylogenetic tree is constructed using MUSCLE for multiple alignment, Neighbour-Joining method and Kimura-2-parameter distance. The numerical figures at each node represent Bootstrap values ($n=5000$). The taxonomic breakdown of candidate genes is based on three fully sequenced shotgun metagenomes of tropical soils in secondary Neotropical Panamanian forests.

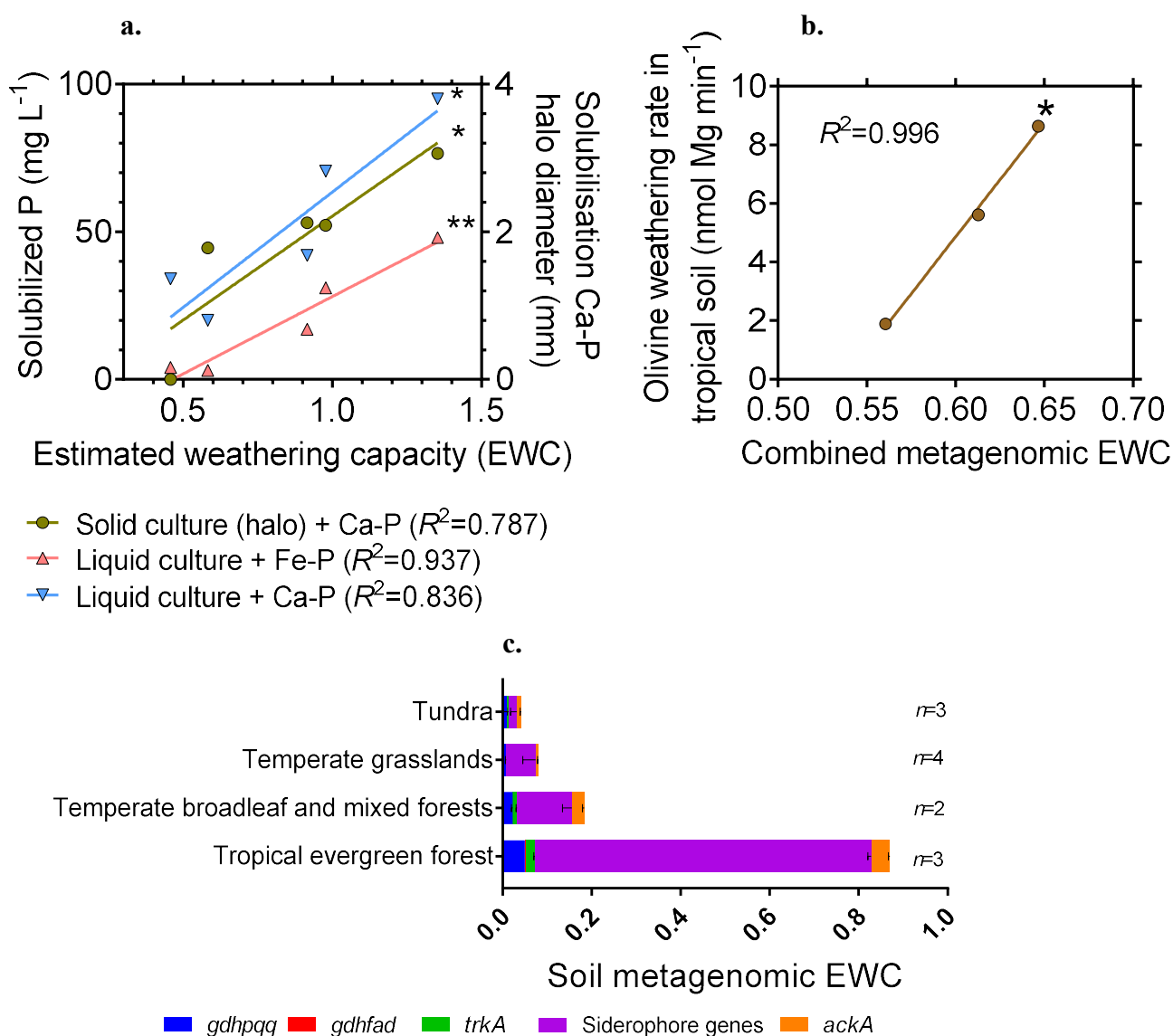


Figure 4. Estimated weathering capacity (EWC) modelled using target weathering genes correlates with weathering rates (a) *in vitro* (by single bacterial species) and (b) *in vivo* (by the combined mineral+soil microbial communities) and can be utilised for whole-biome modelling of microbial weathering activity (c).

ADDED!

Supplementary Table 1. Bacterial genera that have been demonstrated to enhance the process of rock and mineral weathering *in vitro* with supporting references.

Genus	References
<i>Arthrobacter</i>	1
<i>Microbacterium</i>	
<i>Promicromonospora</i>	
<i>Pseudonocardia</i>	
<i>Nocardiopsis</i>	
<i>Dietzia</i>	
<i>Pseudomonas</i>	
<i>Actinopolyspora</i>	2
<i>Actinomadura</i>	
<i>Kitasatospora</i>	
<i>Nocardioides</i>	
<i>Kibdelosporangium</i>	
<i>Streptomyces</i>	3
<i>Cabelleronia</i>	
<i>Pseudomonas</i>	4
<i>Novosphingobium</i>	
<i>Microbacterium</i>	
<i>Acinetobacter</i>	
<i>Streptomyces</i>	
<i>Rhizobium</i>	
<i>Chitinophaga</i>	
<i>Pantoea</i>	
<i>Staphylococcus</i>	
<i>Leclercia</i>	
<i>Ensifer</i>	
<i>Arthrobacter</i>	
<i>Bacillus</i>	
<i>Lysinibacillus</i>	
<i>Comamonas</i>	
<i>Staphylococcus</i>	
<i>Acinetobacter</i>	
<i>Planococcus</i>	
<i>Microbacterium</i>	
<i>Paenibacillus</i>	
<i>Paracoccus</i>	
<i>Sporosarcina</i>	
<i>Ralstonia</i>	
<i>Bacillus</i>	
<i>Burkholderia</i>	
<i>Erwinia</i>	5
<i>Burkholderia</i>	
<i>Rhizobium</i>	
<i>Staphylococcus</i>	
<i>Paenibacillus</i>	

<i>Bacillus</i>	5
<i>Pseudomonas</i>	
<i>Arthrobacter</i>	
<i>Bacillus</i>	
<i>Lysinibacillus</i>	
<i>Rhodococcus</i>	
<i>Massilia</i>	
<i>Serratia</i>	
<i>Stenotrophomonas</i>	
<i>Curtobacterium</i>	
<i>Rhizobium</i>	6
<i>Paenibacillus</i>	7
<i>Bacillus</i>	
<i>Arthrobacter</i>	
<i>Curtobacterium</i>	
<i>Ochrobactrum</i>	
<i>Acinetobacter</i>	
<i>Pantoea</i>	
<i>Pseudomonas</i>	
<i>Flavobacterium</i>	
<i>Ensifer</i>	
<i>Rhizobium</i>	
<i>Pantoea</i>	
<i>Stenotrophomonas</i>	
<i>Bacillus</i>	
<i>Novosphingobium</i>	
<i>Serratia</i>	
<i>Pseudoxanthomonas</i>	
<i>Solibacillus</i>	
<i>Agromyces</i>	
<i>Caulobacter</i>	
<i>Achromobacter</i>	
<i>Enterobacter</i>	
<i>Microbacterium</i>	
<i>Sphingomonas</i>	
<i>Exiguobacterium</i>	
<i>Paenibacillus</i>	
<i>Brachybacterium</i>	
<i>Curtobacterium</i>	
<i>Phyllobacterium</i>	
<i>Massilia</i>	
<i>Granulicatella</i>	
<i>Chryseobacterium</i>	
<i>Vogesella</i>	
<i>Mitsuaria</i>	
<i>Agrobacterium</i>	8
<i>Sinomonas</i>	
<i>Microbacterium</i>	

<i>Cupriavidus</i>	8
<i>Sphingobium</i>	
<i>Flavobacterium</i>	
<i>Variovorax</i>	
<i>Ensifer</i>	
<i>Pseudomonas</i>	
<i>Brevibacillus</i>	
<i>Acinetobacter</i>	
<i>Burkholderia</i>	
<i>Stenotrophomonas</i>	
<i>Bacillus</i>	
<i>Arthrobacter</i>	
<i>Klebsiella</i>	9
<i>Enterobacter</i>	
<i>Pantoea</i>	
<i>Agrobacterium</i>	
<i>Microbacterium</i>	
<i>Burkholderia</i>	
<i>Myroides</i>	
<i>Acidithiobacillus</i>	10
<i>Acetobacter</i>	
<i>Burkholderia</i>	11
<i>Bacillus</i>	
<i>Ralstonia</i>	
<i>Cupriavidus</i>	
<i>Lysinibacillus</i>	
<i>Microbacterium</i>	
<i>Myroides</i>	
<i>Ochrobactrum</i>	
<i>Enterobacter</i>	
<i>Pseudomonas</i>	
<i>Enhydrobacter</i>	
<i>Leifsonia</i>	
<i>Cellulomonas</i>	
<i>Kocuria</i>	
<i>Providencia</i>	
<i>Arthrobacter</i>	
<i>Curtobacterium</i>	
<i>Agrobacterium</i>	12
<i>Aminobacterium</i>	
<i>Azospirillum</i>	
<i>Labrys</i>	
<i>Rhanella</i>	
<i>Rhizobium</i>	
<i>Sphingomonas</i>	
<i>Achromobacter</i>	
<i>Burkholderia</i>	
<i>Collimonas</i>	
<i>Janthinobacterium</i>	

<i>Acinetobacter</i>	12
<i>Azotobacter</i>	
<i>Geobacter</i>	
<i>Acidithiobacillus</i>	
<i>Citrobacter</i>	
<i>Dyella</i>	
<i>Enterobacter</i>	
<i>Frateuria</i>	
<i>Pseudomonas</i>	
<i>Serratia</i>	
<i>Shewanella</i>	
<i>Arthrobacter</i>	
<i>Bacillus</i>	
<i>Mycobacterium</i>	
<i>Paenibacillus</i>	
<i>Staphylococcus</i>	
<i>Streptomyces</i>	
<i>Rhodococcus</i>	13
<i>Janthinobacterium</i>	
<i>Paenibacillus</i>	
<i>Pseudomonas</i>	
<i>Collimonas</i>	
<i>Sphingomonas</i>	
<i>Bacillus</i>	14
<i>Paenibacillus</i>	
<i>Vibrio</i>	
<i>Xanthobacter</i>	
<i>Enterobacter</i>	
<i>Kluyvera</i>	
<i>Pseudomonas</i>	
<i>Chryseomonas</i>	
<i>Bradyrhizobium</i>	15
<i>Rhizobium</i>	
<i>Cupriavidus</i>	
<i>Burkholderia</i>	
<i>Escherichia</i>	16
<i>Arthrobacter</i>	17
<i>Janthinobacterium</i>	
<i>Leifsonia</i>	
<i>Paucibacter</i>	
<i>Polaromonas</i>	
<i>Pseudomonas</i>	
<i>Rhodococcus</i>	

Supplementary references

1. Potysz, A., Grybos, M., Kierczak, J., Guibaud, G., Lens, P.N. and van Hullebusch, E.D., 2016. Bacterially-mediated weathering of crystalline and amorphous Cu-slugs. *Applied geochemistry*, 64, pp.92-106.
2. Abdulla, H., 2009. Bioweathering and biotransformation of granitic rock minerals by actinomycetes. *Microbial ecology*, 58(4), pp.753-761.
3. Uroz, S. and Oger, P., 2017. *Caballeronia mineralivorans* sp. nov., isolated from oak-*Scleroderma citrinum* mycorrhizosphere. *Systematic and applied microbiology*, 40(6), pp.345-351.
4. Wang, Q., Wang, R.R., He, L.Y., Lu, J.J., Huang, Z. and Sheng, X.F., 2014. Changes in weathering effectiveness and community of culturable mineral-weathering bacteria along a soil profile. *Biology and fertility of soils*, 50(7), pp.1025-1034.
5. Wang, Q., Cheng, C., He, L., Huang, Z. and Sheng, X., 2014. Characterization of depth-related changes in bacterial communities involved in mineral weathering along a mineral-rich soil profile. *Geomicrobiology Journal*, 31(5), pp.431-444.
6. Zhao, F., Qiu, G., Huang, Z., He, L. and Sheng, X., 2013. Characterization of *Rhizobium* sp. Q32 isolated from weathered rocks and its role in silicate mineral weathering. *Geomicrobiology Journal*, 30(7), pp.616-622.
7. Huang, J., Sheng, X.F., Xi, J., He, L.Y., Huang, Z., Wang, Q. and Zhang, Z.D., 2014. Depth-related changes in community structure of culturable mineral weathering bacteria and in weathering patterns caused by them along two contrasting soil profiles. *Appl. Environ. Microbiol.*, 80(1), pp.29-42.
8. Zhang, Z., Huang, J., He, L. and Sheng, X., 2016. Distinct weathering ability and populations of culturable mineral-weathering bacteria in the rhizosphere and bulk soils of *Morus alba*. *Geomicrobiology Journal*, 33(1), pp.39-45.
9. Zhang, C. and Kong, F., 2014. Isolation and identification of potassium-solubilizing bacteria from tobacco rhizospheric soil and their effect on tobacco plants. *Applied soil ecology*, 82, pp.18-25.
10. Pollmann, K., Kutschke, S., Matys, S., Kostudis, S., Hopfe, S. and Raff, J., 2016. Novel biotechnological approaches for the recovery of metals from primary and secondary resources. *Minerals*, 6(2), p.54.
11. Yu, Y., Sheng, X., He, L. and Huang, Z., 2016. Linkage between culturable mineral-weathering bacteria and their weathering effectiveness along a soil profile. *Geomicrobiology Journal*, 33(1), pp.10-19.
12. Uroz, S., Calvaruso, C., Turpault, M.P. and Frey-Klett, P., 2009. Mineral weathering by bacteria: ecology, actors and mechanisms. *Trends in microbiology*, 17(8), pp.378-387.
13. Lapanje, A., Wimmersberger, C., Furrer, G., Brunner, I. and Frey, B., 2012. Pattern of elemental release during the granite dissolution can be changed by aerobic heterotrophic bacterial strains isolated from Damma glacier (Central Alps) deglaciated granite sand. *Microbial ecology*, 63(4), pp.865-882.
14. Vazquez, P., Holguin, G., Puente, M.E., Lopez-Cortes, A. and Bashan, Y., 2000. Phosphate-solubilizing microorganisms associated with the rhizosphere of mangroves in a semiarid coastal lagoon. *Biology and Fertility of Soils*, 30(5-6), pp.460-468.
15. Marra, L.M., Oliveira, S.M.D., Soares, C.R.F.S. and Moreira, F.M.D.S., 2011. Solubilisation of inorganic phosphates by inoculant strains from tropical legumes. *Scientia Agricola*, 68(5), pp.603-609.
16. Garcia, B., Lemelle, L., Rose-Koga, E.F., Telouk, P., Gillet, P. and Albarede, F., 2005, December. The Mg-isotope biosignature of *Escherichia coli*-mediated olivine dissolution. In *AGU Fall Meeting Abstracts*.

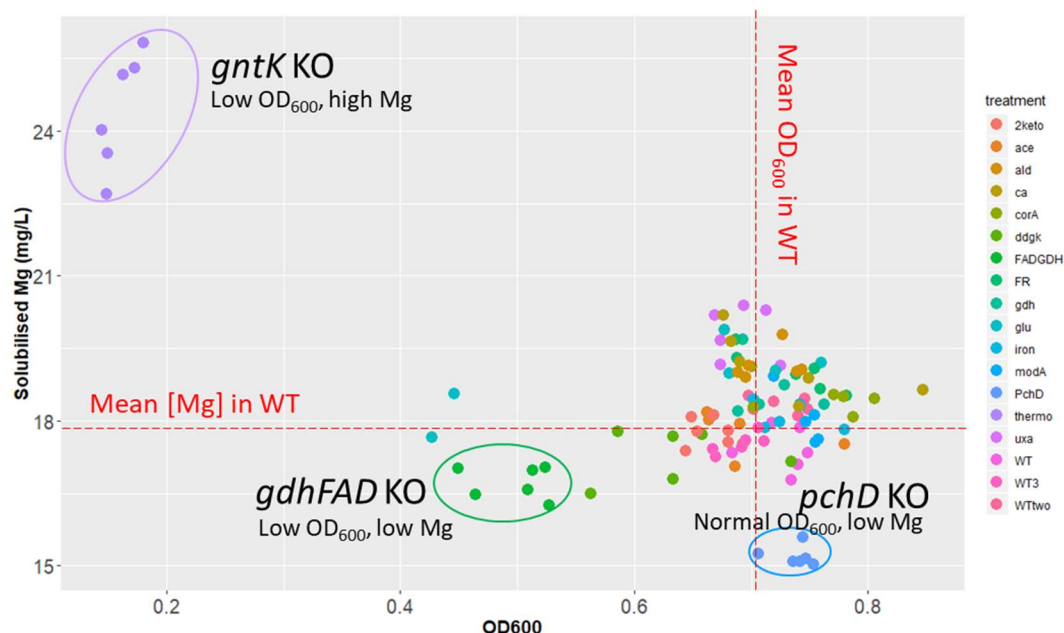
17. Frey, B., Rieder, S.R., Brunner, I., Plötze, M., Koetzsch, S., Lapanje, A., Brandl, H. and Furrer, G., 2010. Weathering-associated bacteria from the Damma glacier forefield: physiological capabilities and impact on granite dissolution. *Appl. Environ. Microbiol.*, 76(14), pp.4788-4796.

ADDED!

Supplementary Note 1. The effect of bacterial culture density as measured by OD₆₀₀ on *in vitro* olivine weathering rate

The optical density (OD) of a bacterial culture measured at $\lambda=600$ nm (hereafter referred to as OD₆₀₀) provides a reliable estimate of bacterial growth and number of cells for a particular species. In our study, the number of cells in turn is expected to affect mineral weathering, as more cells and greater metabolic activity may translate to additional output of weathering agents such as organic acids and siderophores and enhancement in bacteria-to-mineral surface interactions.

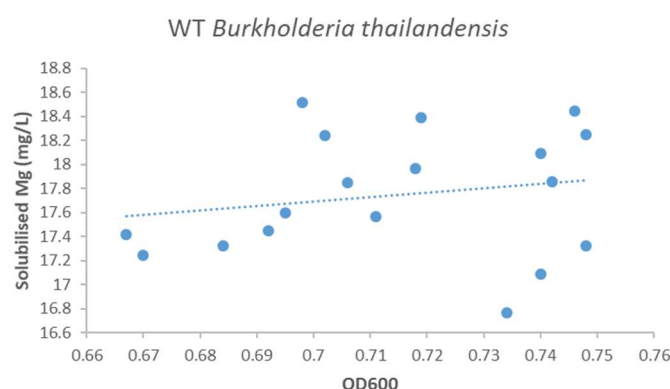
To test for this hypothesis, we correlated the concentration of solubilised Mg in cultures by the OD₆₀₀ readings across all single gene KO mutants and wild type *B. thalaindensis* cultures in our experiments. KO mutants exhibited non-parametric distribution and Spearman correlation test was applied in place of Pearson, revealing no significant correlation between OD₆₀₀ and solubilised Mg (Spearman test $P>0.10$). The three single gene mutant backgrounds with largest effect sizes in terms of weathering – *gntK*⁻, *gdhFAD*⁻ and *pchD*⁻ also confirmed that pattern showing that both normal and low OD₆₀₀ can co-occur with low levels of weathering (as measured by solubilised Mg concentration in *gdhFAD*⁻ and *pchD*⁻) and that low OD₆₀₀ can co-occur with both high weathering (as in *gntK*⁻) and low weathering (as in *gdhFAD*⁻).



As already outlined in the Main Text, these observations can be explained by differences in the metabolism of these mutants. Knock-out *gdhFAD*⁻ mutants are interrupted in their ability to oxidise glucose to gluconic acid thus unable to undergo their preferred mode of glucose assimilation through the semi-phosphorylative ED pathway. This may be the cause for a lag

in their exponential phase and lower OD₆₀₀ readings until they shift their expression from the gluconic acid-producing semi-phosphorylative ED pathway to alternative routes of glucose catabolism such as the classic ED or EMP glycolysis. As a result, less gluconic acid is formed and weathering rates decline. In contrast, *gntK*⁻ mutants will be able to oxidise the glucose (the only C source in our specially designed medium) to gluconic acid but will not be able to transport it inside the cytosol and metabolise it. Consequently, this causes the build-up of large unused and non-utilisable amounts of C substrate as gluconic acid in the medium stunting growth and triggering proportional increases in olivine weathering. On the other hand, *pchD*⁻ mutants incapable of producing siderophores will be able to normally assimilate glucose and reveal normal OD₆₀₀ readings. This indicates that although the lack of siderophores causes a substantial decline in mineral weathering, sufficient levels of Fe and other micronutrients must be generated by alternative weathering mechanisms (such as gluconic acid production) as to support bacterial growth.

The variation in growth in WT Bt bacteria also does not correlate with their ability to solubilise Mg from olivine (Pearson test, $P > 0.10$) confirming once again that the effects of bacteria on weathering are due to qualitative (differences in metabolism) than purely quantitative (number of cells) factors.



Supplementary Note 2. The amount of gluconic acid in culture filtrate from WT and *gntK* mutants

	<i>gntK</i> KO	WT	<i>gdhFAD</i> KO
Gluconic acid* (mM)	3.02 (0.29)**	not detected	not detected
pH	3.70 (0.03)***	4.40 (0.04)	4.65 (0.02)

*Gluconic acid was determined in syringe 0.22 μ m filtrates of 10-day old cultures using the Megazyme D-gluconic acid/D-glucono- δ -lactone kit.

**The number in brackets represent S.E.M. $n = 3$ replicates in the case of both *gntK*, WT and *gdhFAD* cultures.

***In the case of pH $n = 6$.

Notes: These results support the view that *gntK* mutants are unable to transport gluconic acid inside the cell causing the build up of gluconic acid and acidity in the medium causing futile cycles of glucose oxidation to gluconic acid without assimilation. In cells of WT bacteria gluconic acid is transported inside the cell, the cell assimilates it and gluconic acid does not build up in the medium generating less acidity. In *gdhFAD* cells glucose is acquired and assimilated directly without prior oxidation to gluconic acid thus eliminating the H⁺ generation step of glucose oxidation resulting in higher solution pH. It is somewhat puzzling that WT filtrates did not contain even slight traces of gluconic acid. Previous work has identified that gluconic acid accumulation in the solution is temporary specific, dependent upon the growth stage of bacterial cultures with more gluconic acid detected in the first couple of days of growth¹. As such, future work should use more time points to establish the gluconic acid accumulation in WT BtE264 bacteria.

1. Chen, W., Luo, L., He, L.Y., Wang, Q. and Sheng, X.F., 2016. Distinct mineral weathering behaviors of the novel mineral-weathering strains *Rhizobium yantingense* H66 and *Rhizobium etli* CFN42. *Appl. Environ. Microbiol.*, 82(14), pp.4090-4099.

Discussion and Conclusions

Discussion and Conclusions

The biogeochemistry and belowground metagenomics of N₂-fixers: in search of fixer effects beyond nitrogen enrichment

Dimitar Z. Epihov^{1*} and David J. Beerling¹

¹*Department of Animal and Plant Sciences, University of Sheffield, Sheffield S10 2TN, UK*

*for correspondence

Abstract

The enrichment in ecosystem N stocks caused by the presence of diverse N₂-fixing trees in successional settings is well-documented. However, other biogeochemical effects brought up by N₂-fixers in such systems have been largely overlooked. Here, we summarise data from several field studies and show that N₂-fixing trees cause a significant 52-60% increase in rock weathering rates accompanied by a release of key mineral nutrients. Utilising metagenomic and belowground microbial community data we show that these weathering enhancements are linked to substantial enrichments in microbial respiration, lithotrophy and gluconic acid production beneath N₂-fixers likely driven by high inputs of fixed N as well as symbiotic attractions. Consequently, we argue that the ecosystem services provided by N₂-fixing trees during successional and agroforestry scenarios are more far-reaching than previously realised.

Introduction

Symbiotic N₂-fixing angiosperms are members of a diverse monophyletic clade within the rosids known as the Nitrogen Fixing Clade (NFC)¹. Within the clade, only members of the Leguminosae (Fabaceae) and Cannabaceae (e.g. *Parasponia*) form symbiotic root nodules harbouring rhizobia (including α and β -proteobacteria), whereas N₂-fixing members of the Betulaceae, Rosaceae, Rhamnaceae, Casuarinaceae among others are defined as actinorrhizal, forming root nodules sheltering members of the actinomycete bacterial genus of *Frankia*^{1,2}. Within the NFC, the legume family is the most speciose, with very diverse geographic

distribution. Legume trees are prominent elements of many tropical forests worldwide³. Legume trees are also frequently planted by man with multiple examples from agroforestry systems^{4,5,6}, urban parks and green areas⁷, reforestation projects and disturbed and degraded land management⁸. Highly praised factors contributing to their widespread presence in such systems are their characteristically fast growth, capacity to obtain atmospheric N₂ through nodulating symbiotic partners and ability to perform well under unfavourable edaphic conditions. The same qualities may also account for some of legume's prodigious capability to invade natural systems^{9,10,11}.

Symbiotic N₂-fixers are also particularly important during ecosystem succession. Examples of highly abundant N₂-fixers (leguminous and actinorrhizal plants) are well documented during both primary succession (**Table 1**) and secondary succession, particularly under N-limited conditions (**Table 2; Figure 1 and 2**). Under these conditions, a commonly described phenomenon is that N₂-fixers provide high amounts of fixed N enriching the soil and the aboveground biomass in N through redistribution of their N-rich litter (**Tables 1 & 2**). Here, we argue that N₂-fixers modify ecosystem biogeochemical cycling beyond their effect on N stocks including weathering.

In our previous work, we have hypothesized that the evolution of legume-rich tropical forests in the early Cenozoic (58-42 Mya) triggered increases in silicate rock weathering (Introduction) via cascading effects of fixer-mediated increase in ecosystem N stocks. Here, we conclude our investigations into tropical legume trees and their soil-mineral microbiome in respect to biogeochemistry and routes to enhanced mineral access by comparing our findings to other studies and highlighting common fixer-effects in the biogeochemical cycling of fixer-rich ecosystems.

1. Symbiotic N₂-fixation increases weathering inputs during succession

Our field experiments demonstrate that forest native N₂-fixing legumes drive significantly greater silicate weathering rates in their rooting zones than non-fixing non-legume trees in both Panamanian (Chapter 1) and Australian (Chapter 2) post-abandonment secondary tropical forests (**Figure 3a,b**). Interestingly, this ability of legume trees coincides with legume-specific reduction in pH (**Figure 3e,f**) and peak in nodule biomass during secondary succession (**Figure 3i,j**). Although, our studies represent the only large-scale comparisons

specifically designed to test the differences in *in situ* silicate weathering between early successional tropical N₂-fixers and non-fixers, other studies investigating the effects of temperate N₂-fixing trees on soil and nutrient stocks during successional dynamics are scarce but nevertheless corroborate our findings. For instance, a study of the actinorrhizal *Alnus rubra* of secondary forests in USA shows substantially higher removal rates of Ca, Mg and K contributed to weathering of primary minerals (weathering fluxes) beneath the early successional *Alnus* stands relative to nearby mid-successional stands of the evergreen conifer *Pseudotsuga menziesii* on the same soil (**Figure 3c**). The authors also showed that weathering fluxes beneath *A. rubra* were among the highest recorded relative to other stands of deciduous angiosperm trees as diverse as *Liriodendron tulipifera*, *Quercus montana* and a mixed *Quercus-Carya* system (**Figure 3c**). Similarly, another study investigating primary succession following glacier recession in Alaska, USA reveals the highest rates of calcium carbonate (calcite) mineral weathering in soil beneath the N₂-fixing actinorrhizal *Alnus crispa* and *Dryas* spp. relative to other early successional non-fixing trees such as *Populus trichocarpa* and *Salix* spp. (**Figure 3d**). Last but not least, plant-free soil mesocosms containing fermentative F layer soil from either an alder or a conifer stand incubated with phosphate rock for 30 days revealed that alder soil mesocosms drove 1.8-fold greater mineralization of PO₄³⁻ from phosphate rock than did the conifer F layer soil mesocosms¹² with A layer soil showing no significant difference between conifer and alder soil treatments¹².

A meta-analysis of weathering rates was performed using 8 studies of field experiments in which N₂-fixer trees and non-fixers were tested simultaneously. Three of the studies included basalt bags deposited in topsoil in three biomes in Australia: tropical forest (Chapter 2), forest-savanna transition (unpublished author data) and savanna (unpublished author data). Likewise, the same studies were repeated using dunite. Another study was from secondary tropical forests in Panama (Chapter 1) and last was the study measuring carbonate weathering in the tundra in Alaska beneath primary successional species. The red alder (*Alnus rubra*)-Douglas fir (*Pseudotsuga menziessii*) study was not used as it measures weathering fluxes rather than weathering rates. For each study, weathering rates were equilibrated to the mean and measured in fold relative to the mean of the study (sample/mean). Data was categorised in tree species for both N₂-fixers and non-fixers. Data are summarised in **Figure 4**. Both paired t-test and unpaired Mann-Whitney tests unanimously revealed 52-60% significantly greater weathering rates in N₂-fixers than non-fixers (**Figure 4**).

In addition, to greater weathering, another shared feature beneath successional leguminous and actinorrhizal N₂-fixing trees is their more acidic soil (**Figure 3e,f,g,h**). The early successional peak in nodule biomass typical to pioneering legume trees (**Figure 3i,j,l**), coinciding with the highest inputs of N, also co-occurs with the highest recorded differences in weathering rates between N₂-fixers and non-fixers (**Figure 3b,f,j**).

Fixer-mediated enhancement in weathering is likely to increase nutrient availability to plants and soil solution concentrations in the ecosystem. In support, we find that basalt from beneath *Acacia celsa* in Australian secondary forests sustained consistently greater P and K leaching rates throughout the secondary forest chronosequence (Chapter 2). Molybdenum (Mo) leaching from dunite was greatest in 20-year old sites where nodule biomass in soil and bradyrhizobial *nifK* (Fe-Mo containing nitrogenase) in weathered basalt were at their highest (Chapter 2). These observations are consistent with our hypothesis that high weathering traits evolved in legumes in response to the actively P and Mo demanding N₂-fixation symbioses¹³. In the *Alnus rubra*-dominated systems that exhibited higher mineral weathering than *Pseudotsuga menziesii*, soil solution concentrations of Ca, Mg, K were also greater for all three soil horizons and leaf concentrations of Mg and K were significantly greater in *Alnus* than in *Pseudotsuga*¹⁴. Similarly, limed soils beneath *Robinia pseudoacacia* exhibited significantly greater soil solution concentrations of Ca than the limed soils beneath pine-hardwood mixed stands¹⁵ (**Figure 3k**). In another study, greater available K, Ca and P were observed in soil beneath *Robinia pseudoacacia* relative to a mixed *Quercus-Pinus* system on the same young soils with authors advocating increased primary mineral weathering as the likely driver¹⁶.

On geologically young soils, apatite mineral weathering is the most important input source of P in the system¹⁷. Under such conditions, enhanced weathering of primary minerals by N₂-fixers may account for some boost in P availability to such fixing legume trees. Examples typifying this scenario would include forests on geologically young soils with high mineral content such as alfisols and cambisols dominating most of the forested landscapes in temperate areas or on soils resulting from recent volcanic activity such as the ones beneath some tropical forests in Hawaii¹⁸.

However, many tropical soils, including the highly weathered oxisols from our Australian and Panamanian field sites are geologically old and contain little primary minerals with most P occluded in complexes with secondary iron and aluminium minerals^{19,17}. Our observation that

legume trees promote the dissolution of primary silicates in such soils is likely transferable to secondary minerals as dissolution of secondary minerals follows the same set of rules as that of primary minerals – with the concentration of H^+/OH^- , organic acids and CO_2 , as well as precipitation and temperature playing important roles^{20,21}. For instance, total soil concentrations of Al (mainly contained within kaolinitic secondary minerals) was significantly lower beneath N_2 -fixing trees in both of our tropical forest studies in Panama and Australia indicating increased weathering and leaching of kaolinite similar to lower exchangeable Al beneath *Robinia pseudoacacia* in temperate successional forests¹⁶. Therefore, it seems likely that the link between N_2 -fixers and high weathering rates would be valuable not only on geologically young but also geologically old soils. Consequently, it is conceivable that N_2 -fixers during successional scenarios do not only facilitate ecosystem development through their inputs of fixed N but may also supply greater availability of K, Ca, Mg, P and Mo plundered from their higher weathering yields.

2. Metagenomic and biogeochemical evidence for the role of fixed N inputs, nitrification and nitrate leaching in weathering through cation removal and acidification

Increased inputs of N resulting from N_2 -fixers can modify the biogeochemistry of an ecosystem resulting in (1) decline in soil C:N ratios, (2) increases in soil ammonia/nitrate, (3) increased nitrification rates and (4) elevated nitrate leaching. Ultimately, these processes drive acidification and removal of counterbalancing cations (Mg, Ca, K) which can stimulate weathering through acidolysis and through shifting the equilibrium to further dissolution, respectively.

In the microbial community increased inputs of fixed N can be expected to trigger: (1) rise in N cycling microorganisms and (2) genes involved in microbial nitrification.

Biogeochemical evidence from tropical forest legumes in Panama indicates that soil beneath N_2 -fixers contains lower C:N ratio characteristic to high N inputs. Non-fixing trees near N_2 -fixers exhibit similarly low soil C:N ratio unlike non-fixers far from fixers with significantly higher C:N and receiving less inputs of fixed N. Similarly, lower soil C:N ratios are also observed in temperate N_2 -fixing trees such as the legume *Robinia pseudoacacia* and in actinorrhizal trees (**Figure 5a-c**). In support of the role of N inputs in driving weathering, the soil C:N ratio significantly correlated with weathering rates (**Figure 5a**) in the study of

weathering rates in Panamanian tropical forests. High N inputs increase nitrification that is the microbial transformation of NH_3 to NO_3^- which generates acidity and promotes cation and nitrate leaching – all factors understood to positively affect rock weathering.

Meta-analysis of community sequencing-based profiling of soil beneath N_2 -fixing plants (including Panamanian tropical legume trees, pigeon pea *Cajanus cajan*, *Mimosa debilis* and *Alnus* spp.) reveal greater relative abundance of N-cycling microorganisms in soil of N_2 -fixers than in non-fixers within the same study (Mann-Whitney test, $P < 0.05$; **Figure 5e,f**). This is consistent with higher N inputs causing enrichment in N cycling microbial taxa in soil (**Figure 5d**). Meta-analysis coupling metagenomics of the mineral-associated communities of soil-deposited rock material to weathering rates indicates that the abundance of all four major nitrification genes (ammonia monooxygenase *AMO*, hydroxylamine oxidase *HAO*, nitrate reductase alpha subunit *narG* and nitrate reductase beta subunit *narH*) positively associates with silicate weathering rates (**Figure 5g**). These significant correlations between nitrification genes and weathering rates of deposited fresh silicate rock material complements biogeochemical evidence of greater nitrification rates and weathering fluxes beneath N_2 -fixing trees of temperate areas such as *Alnus rubra* and *Robinia pseudoacacia*.

3. Higher microbial respiration in stands rich in N_2 -fixing trees and its role in weathering

The high N litter of many N_2 -fixers decomposes faster^{22,23}. In terms of microbial metabolism, fast decomposition translates into high respiration rates. A field study comparing the effects of invasive N_2 -fixers revealed significantly greater basal respiration rates per unit soil organic matter beneath a recent invasion by N_2 -fixing *Acacia longifolia* than the native non-fixing vegetation in sand dune ecosystems in Portugal²⁴. Similarly, microbial respiration per unit soil organic matter beneath a stand planted to the N_2 -fixing legume tree *Robinia pseudoacacia* was 4-fold higher than that in the native secondary oak (*Quercus liaotungensis*) forest in the Loess Plateau in China²⁵. Secondary forests in the Czech Republic dominated by the actinorrhizal N_2 -fixing trees *Alnus glutinosa* and *Alnus incana* revealed ~3- and 11-fold higher basal respiration rates than *Quercus rubra* and *Pinus sylvestris*-dominated forests, respectively²⁶. To cross-validate the field observations of higher respiration rates beneath stands dominated by N_2 -fixing trees of temperate and Mediterranean latitudes to our field

trials in tropical forests, we utilized the large collection of sequenced metagenomes of soil-mineral samples we have previously characterised and generated.

Unfortunately, tropical forests are rarely as clear-cut as temperate forests in terms of dominant species and pure legume or non-legume natural stands are hard to come by. Our previous work clearly illustrates this problem: our community profiling in Panama revealed that non-fixing trees neighbouring N₂-fixing legumes exhibit lower soil C:N ratios and microbial communities more similar to those of N₂-fixers than to those of non-fixers far from N₂-fixers (Chapter 1). To circumvent this obstacle, we grouped our 46 metagenomic samples into highly fixing, highly nodulated forest samples in which N-rich litter is actively re-distributed between fixers and non-fixers and in forests of low N₂-fixation and nodulation rates. The first group (“high-fixing forest samples”) contained soil-mineral metagenomic samples from fixers ($n=6$ dunite) and non-fixers near fixers ($n=3$ dunite) from highly nodulated 17-year old forests in Panama as well as soil-mineral samples from fixing ($n=4$ for dunite and 4 for basalt) and neighbouring non-fixing trees ($n=2$ for dunite and 2 for basalt) in the most nodulated 20-year old forest site in Australia. In contrast, the second group (“low-fixing forest samples”) consisted of soil-mineral metagenomic samples from non-fixers far from fixers ($n=3$ dunite) from our tropical forests sites in Panama and samples from the 12.5 and 48-year old Australian forests where assessment of nodulation rates showed low to near-absent symbiotic N₂-fixation ($n=14$ for dunite and 14 for basalt). Because all fresh mineral samples are microbially-speaking identical prior their deposition in the 0-10cm topsoil, they represent a convenient method to assess the short-term effects driven by the tree and its rhizosphere during the 8-10 months weathering period in soil.

High-level metagenomic comparisons revealed that the abundance of genes involved in microbial respiration was consistently greater (Mann-Whitney test, $P<0.001$) in metagenomes from high-fixing forests ($n=21$) than in metagenomes from low-fixing forests ($n=25$) – a trend consistent across soil-deposited rock types (dunite and basalt) and forest locations (Panama and Australia; **Figure 6a**). The sum of microbial respiration genes also consistently positively correlated with silicate weathering rates (**Figure 6a**). Microbial respiration is a major source of CO₂ in soil and its associated weathering-potent carbonic acid which may be one reason for their strong association in addition to glycolytic (gluconic, acetic) and Krebs (citric, oxalic, malic) organic acids which are also produced by actively respiring microorganisms. In addition to providing N necessary for respiration (otherwise N limitation puts a break on the

Krebs cycle and glycolysis²⁷), fast decomposition would also be expected to quickly release labile C such as starch from plant materials explaining the strong correlation between the gene abundance of microbial glucoamylase and weathering and its higher abundance in high-fixing forests (not shown).

4. N₂-fixing trees are linked to more abundant belowground lithotrophy

In essence, microbial respiration generates energy from coupling electron transfer from a donor to an acceptor across a respiratory membrane chain, generating a proton motive force used for synthesis of ATP. Many microorganisms use organic molecules for donors of electrons (chemoorganotrophs) and fast-decomposing legume litter may provide large quantities of labile C sources quickly assimilated by microorganisms as discussed above. However, other microorganisms utilise inorganics as electron donors (chemolithotrophs) including sulphate oxidisers, ammonia oxidisers, iron oxidisers, hydrogen oxidisers and some sulfate reducers.

A clearly established example for the effects of N₂-fixing trees on lithotrophic communities is the oxidation of ammonia and nitrite during nitrification. By supplying N-rich litter, high loads of NH₃ released through ammonification feeds into increases in nitrifying lithotrophic organisms and nitrification rates beneath such fixer-rich forests as we have already discussed. Indeed, high-fixing forest metagenomic samples were significantly enriched in nitrification genes showing 1.3-fold greater *narG* and *narH*, 2.2-fold greater *HAO* and 3.7-fold greater *AMO* gene abundances than those of low-fixing forests (Mann-Whitney test, $P < 0.01$, except for *HAO* where $P < 0.001$; **Figure 6d**). However, it is comparatively less understood how N₂-fixers impact other lithotrophic groups in soil.

Our soil-mineral metagenomes from high-fixing forests are significantly enriched in the sulfur (S) oxidation pathway relative to low-fixing metagenomes (+1.1-fold; Mann-Whitney test, $P < 0.01$; **Figure 6b**) in both Panamanian and Australian tropical forests and in both soil-deposited dunite and basalt rocks. Key S oxidation genes significantly enriched in high-fixing forests include sulfite oxidase (1.1-fold, Mann-Whitney test, $P < 0.05$), sulfite dehydrogenase *SoxD* (1.2-fold, Mann-Whitney test, $P < 0.01$) and sulfane dehydrogenase subunit *SoxC* cytochrome subunit (1.2-fold, Mann-Whitney test, $P < 0.01$). Previous molecular studies have linked the number of *Sox* genes in soil communities to measured sulfur oxidation

rates²⁸ reaffirming that *Sox* gene abundance can reliably predict *in situ* S oxidizing activity. In addition, 16S rRNA-based community analyses using the SSU SILVA database revealed enrichment of several S oxidizing genera (*Thioalkalivibrio*, *Allochromatium*, *Thiobacter*, sulfur-oxidiser OBII5 – **Chapter 1: Figure 3b**) in soil microbiomes beneath N₂-fixers.

The S oxidation gene enrichment in high-fixing forests (1.1-1.2-fold) is more modest than that of their respective lithotrophic nitrification gene analogues which exhibited 1.3-3.7-fold increase in high-fixing forest soil-mineral metagenomes than in those of low-fixing forests (**Figure 6d**). Meta-analysis of the RAINFOR foliar dataset available from the TRY database shows that N₂-fixing legume trees of tropical forests do not have significantly different foliar S levels than non-fixing trees (**Chapter 1: Supplementary Figure 4**). Those findings suggest that unlike the direct well established link between N inputs of N₂-fixers and increased nitrification, the link between N₂-fixers and increased S oxidation may be indirect and unlinked to higher S inputs.

One possible cause for increased S oxidation potential beneath actively fixing legume trees is the coupling between nitrate reduction and S oxidation. Previously, nitrate addition has been demonstrated to increase oxidation of H₂S to SO₄²⁻ in wetland ecosystems as various microorganisms can couple SO₄²⁻ oxidation to denitrification (NO₃⁻ to N₂) or dissimilatory nitrate reduction (NO₃⁻ to NH₃)²⁹. In support, dissimilatory nitrite reductases and denitrification, both pathways that can be physiologically coupled to S oxidation, showed 2.0-fold and 1.2-fold higher abundance in high-fixing than low-fixing forest samples (**Figure 6e**; denitrification pathway not shown), respectively (Mann-Whitney test, P<0.01 and P<0.05). However, further studies are advised before any definitive mechanisms for the recorded link between S oxidation and high symbiotic N₂-fixation are established.

Metagenomic samples from high-fixing forests also contained significantly more hydrogenases than non-fixing forests (**Figure 6c**; Mann-Whitney test, P<0.001). We show that a total of 26 hydrogenase-encoding genes exhibited a range of 1.2-2.4-fold significant enrichment (Mann-Whitney test, 8 genes P<0.001, 7 genes P<0.01, 11 genes P<0.05) in high-fixing relative to low-fixing forest soil-mineral metagenomes. Hydrogenases are the main enzymes catalysing the lithotrophic H₂ oxidation reaction which generates energy by oxidizing H₂ gas to H₂O and electrons and subsequent electron transfer across the membrane respiratory chain. N₂-fixing legume trees may drive the observed enrichment in H₂ oxidizing metabolism by generating large quantities of H₂ gas as a by-product of their nitrogen fixation

reaction producing a molecule of H₂ per molecule of N₂ fixed. This phenomenon of beneficial effects to plant growth³⁰ has been recorded for legume crops with studies showing greater H₂ gas concentration in legume rhizospheres³⁰. Interestingly, hydrogenase abundance in the metagenome also strongly correlated with weathering rates (**Figure 6c**) suggesting an important role for H₂ oxidative metabolism in bacteria-rock interactions.

High respiration rates beneath N₂-fixers may cause the depletion of O₂ locally and consequent formation of anaerobic microsites at certain soil loci. Such loci would be rich in anaerobic food chain organisms including sulfate reducers, fermenters, methylotrophs and methanogens³¹. Sulfate-reducing lineages use sulfate as their terminal electron acceptor for anaerobic respiration generating H₂S in the process. The metagenomic pathway containing the genes encoding the sulfate reduction-associated complexes shows 1.5-fold significant enrichment in high-fixing metagenomes than low-fixing forest samples (**Figure 6e**; Mann-Whitney test, P<0.05). This finding is further complemented by enrichment of sulfate reducing lineages in the soil microbial communities beneath N₂-fixers including the genera *Desulfococcus*, *Dethiosulfovibrio*, *Desulfobulbus*-like and *Alkalispirillum* (**Chapter 1: Figure 3b**). The metagenomic pathway containing the genes involved in fermentation is also enriched in high-fixing metagenomes relative to low-fixing ones (Mann-Whitney test, P<0.01; **Figure 6e**). Indeed, the 16S rRNA-based reconstruction of the microbial community of N₂-fixers soil is similarly enriched in anaerobic fermentative lineages including many clostridia (not shown). Methylotrophs utilize the gene methanol dehydrogenase to assimilate and convert the C1 compound methanol to formaldehyde with this gene abundance greater in high-fixing than low-fixing metagenomes (Mann-Whitney test, P<0.001; not shown). Combined these findings support the enrichment of anaerobic metabolic niches at localized loci beneath N₂-fixers. They also paint a picture of multiple possible syntrophic (“cross-feeding”) interactions.

For instance, H₂ generation by legume symbioses or by fermenting bacteria can provide H₂ for both hydrogen-oxidizing and lithotrophic sulfate-reducers. Other sulfate-reducers utilise organic electron donors such as acetate, propionate, butyrate and lactate which again may be provided by the activity of fermenting lineages. H₂S produced by sulfate reducers can then be fuelled into sulfur oxidation either alone or coupled to dissimilatory nitrate reduction or denitrification.

5. Symbiotic shaping of belowground microbiomes may stimulate weathering via gluconic and 2-ketogluconic acid production

Our *in vitro* studies with *Burkholderia thailandensis* mutants revealed that gluconic and 2-ketogluconic acids biosynthesis in the Entner-Doudoroff pathway is key in bacteria-driven rock weathering (Chapter 3). However, under field conditions, many other community-driven processes such as nitrification, sulfur oxidation, fermentation etc. occur simultaneously with glucose oxidation making it difficult to isolate the impact that microbial production of gluconic acids will have on rock weathering in the field. We find that the gene encoding the PQQ-dependent glucose dehydrogenase (*gdhPQQ*), the enzyme product of which catalyses the oxidation of glucose to gluconic acid, correlates positively with weathering at slightly lower r-value (Pearson test $P < 0.05$, $r = 0.37$; **Figure 7a**) than processes such as nitrification and sulfur oxidation (**Figures 5 and 6**). The *gdhPQQ* gene is also significantly more abundant in soil-mineral metagenomes beneath high-fixing than non-fixing forests (Mann-Whitney test, $P < 0.01$; **Figure 7a**). It also clearly correlates with other PQQ-containing genes such as PQQ-dependent alcohol dehydrogenase and PQQ-dependent methanol dehydrogenase (**Figure 7b**), suggesting that those enzymes are all under the control of co-enzyme PQQ bioavailability. Indeed, the gene abundance of the pathway encoding co-enzyme PQQ biosynthesis correlated positively with *gdhPQQ* signifying that relationship between PQQ supply and PQQ-dependent enzymes (**Figure 7b**).

Our taxonomic analysis of the *gdhPQQ* gene pool in soil (Chapter 3) revealed that *Bradyrhizobium* contributes 7.2% of all *gdhPQQ* copies in soil which is substantial but somewhat dwarfed by its 40.7% contribution to all gluconate 2-dehydrogenase/*trkA* gene copies in soil which converts gluconic acid to 2-ketogluconic acid, another equally potent weathering agent. Indeed, *trkA* exhibits 1.3-fold greater abundance in high-fixing metagenomes than low-fixing ones at astonishingly low Mann-Whitney P value of 9.06×10^{-5} and it also correlates with field weathering rates (Pearson test $P < 0.01$, $r = 0.41$; **Figure 7a**). Given that *Bradyrhizobium* is the likely nodulating agent of *Acacia celsa* and that it is enriched in basalt communities from beneath *A. celsa* (demonstrated using the *nifK* gene as a marker; see **Chapter 3: Figure 4**) in Australian tropical forests and in weathered soil-dunite (demonstrated using the 16S rRNA gene as a marker; not shown) from beneath N_2 -fixers in Panama, it is very likely that symbiotic attraction may be a driver for the observed enrichment

of *gdhPQQ* and *trkA* genes and indeed enhanced *in situ* production of gluconic acids in high-fixing soil-mineral metagenomes.

A second parallel mechanism explaining the observed enrichment may be the increased respiration rates beneath N₂-fixers releasing labile C and glucose which then attracts more glucose-oxidizing microorganisms converting glucose to gluconic acid and 2-ketogluconic acid through the combined action of *gdhPQQ* and *trkA*.

6. Concluding remarks

Multiple lines of metagenomic and biogeochemical evidence collected in natural systems support substantial functional differences in the microbial community beneath forests containing actively N₂-fixing trees when compared to non- or low-fixing ones. Meta-analysis show that high fixing forest stands support higher abundance of genes involved in microbial respiration, lithotrophic metabolism as well as an enrichment in gluconic acid-producing symbiotic nodulating lineages all converging in 52-60% significantly greater mineral weathering. Consequently, plant communities rich in N₂-fixers during successional scenarios may benefit not only from increases in nitrogen availability but also from more dynamic biogeochemical cycling releasing previously unavailable mineral nutrients. As a consequence of enhanced weathering, the evolution of the first legume-rich tropical forests in the early Cenozoic are likely to have entailed a considerable impact on weathering-driven drawdown of CO₂ and past climate history.

References

1. Li, H.L., Wang, W., Mortimer, P.E., Li, R.Q., Li, D.Z., Hyde, K.D., Xu, J.C., Soltis, D.E. and Chen, Z.D., 2015. Large-scale phylogenetic analyses reveal multiple gains of actinorhizal nitrogen-fixing symbioses in angiosperms associated with climate change. *Scientific reports*, 5, p.14023.
2. Sprent, J.I., 2009. *Legume nodulation: a global perspective*. John Wiley & Sons.
3. Ter Steege, H., Pitman, N.C., Phillips, O.L., Chave, J., Sabatier, D., Duque, A., Molino, J.F., Prévost, M.F., Spichiger, R., Castellanos, H. and Von Hildebrand, P., 2006. Continental-scale patterns of canopy tree composition and function across Amazonia. *Nature*, 443(7110), p.444.
4. Nichols, J.D. and Carpenter, F.L., 2006. Interplanting *Inga edulis* yields nitrogen benefits to

Terminalia amazonia. Forest Ecology and Management, 233(2-3), pp.344-351.

5. Kumar, B.M., Kumar, S.S. and Fisher, R.F., 1998. Intercropping teak with *Leucaena* increases tree growth and modifies soil characteristics. *Agroforestry Systems*, 42(1), pp.81-89.
6. Bauhus, J., Khanna, P.K. and Menden, N., 2000. Aboveground and belowground interactions in mixed plantations of *Eucalyptus globulus* and *Acacia mearnsii*. *Canadian Journal of Forest Research*, 30(12), pp.1886-1894.
7. Moser, A., Rötzer, T., Pauleit, S. and Pretzsch, H., 2015. Structure and ecosystem services of small-leaved lime (*Tilia cordata* Mill.) and black locust (*Robinia pseudoacacia* L.) in urban environments. *Urban Forestry & Urban Greening*, 14(4), pp.1110-1121.
8. Macedo, M.O., Resende, A.S., Garcia, P.C., Boddey, R.M., Jantalia, C.P., Urquiaga, S., Campello, E.F.C. and Franco, A.A., 2008. Changes in soil C and N stocks and nutrient dynamics 13 years after recovery of degraded land using leguminous nitrogen-fixing trees. *Forest Ecology and Management*, 255(5-6), pp.1516-1524.
9. Benesperi, R., Giuliani, C., Zanetti, S., Gennai, M., Lippi, M.M., Guidi, T., Nascimbene, J. and Foggi, B., 2012. Forest plant diversity is threatened by *Robinia pseudoacacia* (black-locust) invasion. *Biodiversity and Conservation*, 21(14), pp.3555-3568.
10. Leary, J.K., Hue, N.V., Singleton, P.W. and Borthakur, D., 2006. The major features of an infestation by the invasive weed legume gorse (*Ulex europaeus*) on volcanic soils in Hawaii. *Biology and Fertility of Soils*, 42(3), pp.215-223.
11. Jovanovic, N.Z., Israel, S., Tredoux, G., Soltau, L., Le Maitre, D., Rusinga, F., Rozanov, A. and Van der Merwe, N., 2009. Nitrogen dynamics in land cleared of alien vegetation (*Acacia saligna*) and impacts on groundwater at Riverlands Nature Reserve (Western Cape, South Africa). *Water SA*, 35(1).
12. Chen, C.S., 1965. Influence of interplanted and pure stands of red alder (*Alnus rubra* Bong.) on microbial and chemical characteristics of a coastal forest soil in the Douglas-fir region.
13. Epihov, D.Z., Batterman, S.A., Hedin, L.O., Leake, J.R., Smith, L.M. and Beerling, D.J., 2017. N₂-fixing tropical legume evolution: a contributor to enhanced weathering through the Cenozoic?. *Proceedings of the Royal Society B: Biological Sciences*, 284(1860), p.20170370.
14. Homann, P.S., Van Miegroet, H., Cole, D.W. and Wolfe, G.V., 1992. Cation distribution, cycling, and removal from mineral soil in Douglas-fir and red alder forests. *Biogeochemistry*, 16(2), pp.121-150.
15. Johnson, D.W., Swank, W.T. and Vose, J.M., 1995. Effects of liming on soils and streamwaters in a deciduous forest: Comparison of field results and simulations. *Journal of environmental quality*, 24(6), pp.1104-1117.
16. Rice, S.K., Westerman, B. and Federici, R., 2004. Impacts of the exotic, nitrogen-fixing black locust (*Robinia pseudoacacia*) on nitrogen-cycling in a pine-oak ecosystem. *Plant Ecology*, 174(1), pp.97-107.
17. Turner, B.L., Lambers, H., Condron, L.M., Cramer, M.D., Leake, J.R., Richardson, A.E. and Smith, S.E., 2013. Soil microbial biomass and the fate of phosphorus during long-term ecosystem development. *Plant and Soil*, 367(1-2), pp.225-234.
18. Vitousek, P.M. and Farrington, H., 1997. Nutrient limitation and soil development: experimental test of a biogeochemical theory. *Biogeochemistry*, 37(1), pp.63-75.
19. Chacon, N., Silver, W.L., Dubinsky, E.A. and Cusack, D.F., 2006. Iron reduction and soil phosphorus solubilization in humid tropical forests soils: the roles of labile carbon pools and an electron shuttle compound. *Biogeochemistry*, 78(1), pp.67-84.
20. Roncal-Herrero, T. and Oelkers, E.H., 2011. Does variscite control phosphate availability in acidic natural waters? An experimental study of variscite dissolution rates. *Geochimica et Cosmochimica Acta*, 75(2), pp.416-426.
21. Gardner, W.K., Barber, D.A. and Parbery, D.G., 1983. The acquisition of phosphorus by *Lupinus albus* L. *Plant and soil*, 70(1), pp.107-124.

22. Schwendener, C.M., Lehmann, J., Rondon, M., Wandelli, E. and Fernandes, E., 2007. Soil mineral N dynamics beneath mixtures of leaves from legume and fruit trees in Central Amazonian multi-strata agroforests. *Acta Amazonica*, 37(3), pp.313-320.
23. Milcu, A., Partsch, S., Scherber, C., Weisser, W.W. and Scheu, S., 2008. Earthworms and legumes control litter decomposition in a plant diversity gradient. *Ecology*, 89(7), pp.1872-1882.
24. Marchante, E., Kj  ller, A., Struwe, S. and Freitas, H., 2008. Short-and long-term impacts of *Acacia longifolia* invasion on the belowground processes of a Mediterranean coastal dune ecosystem. *Applied Soil Ecology*, 40(2), pp.210-217.
25. Shi, W.Y., Tateno, R., Zhang, J.G., Wang, Y.L., Yamanaka, N. and Du, S., 2011. Response of soil respiration to precipitation during the dry season in two typical forest stands in the forest–grassland transition zone of the Loess Plateau. *Agricultural and forest meteorology*, 151(7), pp.854-863.
26. Šourková, M., Frouz, J., Fettweis, U., Bens, O., H  ttl, R.F. and Šantr  ckov  , H., 2005. Soil development and properties of microbial biomass succession in reclaimed post mining sites near Sokolov (Czech Republic) and near Cottbus (Germany). *Geoderma*, 129(1-2), pp.73-80.
27. Shimizu, K., 2014. Regulation systems of bacteria such as *Escherichia coli* in response to nutrient limitation and environmental stresses. *Metabolites*, 4(1), pp.1-35.
28. Tourn  , M., Maclean, P., Condron, L., O'callaghan, M. and Wakelin, S.A., 2014. Links between sulphur oxidation and sulphur-oxidising bacteria abundance and diversity in soil microcosms based on *soxB* functional gene analysis. *FEMS microbiology ecology*, 88(3), pp.538-549.
29. LeBauer, D.S. and Treseder, K.K., 2008. Nitrogen limitation of net primary productivity in terrestrial ecosystems is globally distributed. *Ecology*, 89(2), pp.371-379.
30. Dong, Z., Wu, L., Kettlewell, B., Caldwell, C.D. and Layzell, D.B., 2003. Hydrogen fertilization of soils—is this a benefit of legumes in rotation?. *Plant, Cell & Environment*, 26(11), pp.1875-1879.
31. Sikora, A., Detman, A., Mielecki, D., Chojnacka, A. and B  lszczyk, M., 2018. Searching for Metabolic Pathways of Anaerobic Digestion: A Useful List of the Key Enzymes. In *Biogas*. IntechOpen.
32. Crocker, R.L. and Major, J., 1955. Soil development in relation to vegetation and surface age at Glacier Bay, Alaska. *The Journal of Ecology*, pp.427-448.
33. Bellingham, P.J., Walker, L.R. and Wardle, D.A., 2001. Differential facilitation by a nitrogen-fixing shrub during primary succession influences relative performance of canopy tree species. *Journal of Ecology*, 89(5), pp.861-875.
34. Menge, D.N. and Hedin, L.O., 2009. Nitrogen fixation in different biogeochemical niches along a 120 000-year chronosequence in New Zealand. *Ecology*, 90(8), pp.2190-2201.
35. Blundon, D.J., MacIsaac, D.A. and Dale, M.R.T., 1993. Nucleation during primary succession in the Canadian Rockies. *Canadian Journal of Botany*, 71(8), pp.1093-1096.
36. Dickson, B.A. and Crocker, R.L., 1953. A chronosequence of soils and vegetation near Mt. Shasta, California: II. The development of the forest floors and the carbon and nitrogen profiles of the soils. *Journal of Soil Science*, 4(2), pp.142-154.
37. Batterman, S.A., Hedin, L.O., Van Breugel, M., Ransijn, J., Craven, D.J. and Hall, J.S., 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature*, 502(7470), p.224.
38. Batterman, S.A., Hedin, L.O., Van Breugel, M., Ransijn, J., Craven, D.J. and Hall, J.S., 2013. Key role of symbiotic dinitrogen fixation in tropical forest secondary succession. *Nature*, 502(7470), p.224.
39. Romero-Duque, L.P., Jaramillo, V.J. and P  rez-Jim  nez, A., 2007. Structure and diversity of secondary tropical dry forests in Mexico, differing in their prior land-use history. *Forest Ecology and Management*, 253(1-3), pp.38-47.
40. Boring, L.R. and Swank, W.T., 1984. The role of black locust (*Robinia pseudoacacia*) in forest succession. *The Journal of Ecology*, pp.749-766.

41. Finegan, B. and Delgado, D., 2000. Structural and floristic heterogeneity in a 30-year-old Costa Rican rain forest restored on pasture through natural secondary succession. *Restoration ecology*, 8(4), pp.380-393.
42. Blanc, L., Maury-Lechon, G. and Pascal, J.P., 2000. Structure, floristic composition and natural regeneration in the forests of Cat Tien National Park, Vietnam: an analysis of the successional trends. *Journal of biogeography*, 27(1), pp.141-157.
43. Aweto, A.O., 1981. Secondary succession and soil fertility restoration in South-Western Nigeria: III. Soil and vegetation interrelationships. *The Journal of Ecology*, pp.957-963.
44. Aweto, A.O., 1981. Secondary succession and soil fertility restoration in south-western Nigeria: II. Soil fertility restoration. *The Journal of Ecology*, pp.609-614.
45. Turner, J., Cole, D.W. and Gessel, S.P., 1976. Mineral nutrient accumulation and cycling in a stand of red alder (*Alnus rubra*). *The Journal of Ecology*, pp.965-974.
46. Youngberg, C.T. and Wollum, A.G., 1976. Nitrogen Accretion in Developing Ceanothus Velutinus Stands 1. *Soil Science Society of America Journal*, 40(1), pp.109-112.
47. Cole, D.W., Compton, J., Van Miegroet, H. and Homann, P., 1991. Changes in soil properties and site productivity caused by red alder. In *Management of Nutrition in Forests under Stress* (pp. 231-246). Springer, Dordrecht.
48. Rhoades, C., Oskarsson, H., Binkley, D. and Stottlemeyer, B., 2001. Alder (*Alnus crispa*) effects on soils in ecosystems of the Agashashok River valley, northwest Alaska. *Ecoscience*, 8(1), pp.89-95.
49. Urbanová, M., Šnajdr, J. and Baldrian, P., 2015. Composition of fungal and bacterial communities in forest litter and soil is largely determined by dominant trees. *Soil Biology and Biochemistry*, 84, pp.53-64.50.
50. Sul, W.J., Asuming-Brempong, S., Wang, Q., Turlousse, D.M., Penton, C.R., Deng, Y., Rodrigues, J.L., Adiku, S.G., Jones, J.W., Zhou, J. and Cole, J.R., 2013. Tropical agricultural land management influences on soil microbial communities through its effect on soil organic carbon. *Soil Biology and Biochemistry*, 65, pp.33-38.
51. Lima, A.B., Cannavan, F.S., Navarrete, A.A., Teixeira, W.G., Kuramae, E.E. and Tsai, S.M., 2015. Amazonian Dark Earth and plant species from the Amazon region contribute to shape rhizosphere bacterial communities. *Microbial ecology*, 69(4), pp.855-866.

Figures and Figure Text

Table 1. Examples highlighting the important role of N₂-fixers in ecosystem primary succession

Primary succession examples	Family	Form	N ₂ -fixer	Geology	Climax community	Location	Effect
<i>Alnus crispa</i>	Betulaceae	tree	yes	moraines	Spruce coniferous forest	USA, N America	• Decline in soil pH • Higher carbonate weathering • N fertilisation Ref ³²
<i>Carmichaelia odorata</i>	Leguminosae	shrub	yes	schists	Cool temperate montane rainforest	New Zealand, Oceania	• N fertilisation effect on climax trees Ref ³³
<i>Coriaria arborea</i>	Coriariaceae	shrub	yes	volcanic rhyolite	Cool temperate montane rainforest	New Zealand, Oceania	• N fertilisation effect on climax trees Ref ³⁴
<i>Astracantha aitosensis</i>	Leguminosae	shrub	yes	andesite and basalt tuffs	Mixed temperate forest	Bulgaria, Europe	• N fertilisation?
<i>Hedysarum boreale</i>	Leguminosae	herb	yes	moraines	Shrub, short woodland	Canada, N America	• Nucleation effects (positively associates with other plants) Ref ³⁵
<i>Dryas spp.</i>	Rosaceae	shrub	yes	moraines	Shrub, short woodland	Canada, N America	• Nucleation effects (positively associates with other plants) Ref ³⁵
<i>Purshia tridentata</i>	Rosaceae	shrub	yes	moraines	Coniferous temperate forest	USA, N America	• N fertilisation? Ref ³⁶

Table 2. Examples highlighting the important role of N₂-fixers in ecosystem secondary succession

Secondary successional species	Family	Form	N ₂ -fixers	Climax community	Location	Effects	Reference
<i>Inga cocleensis</i> , <i>Inga thibaudiana</i> , <i>Swartzia simplex</i> , <i>Abarema barbouriana</i> , <i>Platymiscium dimorphandrum</i>	Leguminosae	trees	yes	Wet Tropical forest	Panama, Mesoamerica	• Decline in soil pH • Increase in ecosystem N stocks • Weathering enhancement • Early successional peak in nodulation	Chapter 1 and Ref ³⁷
<i>Acacia celsa</i>	Leguminosae	tree	yes	Wet Tropical forest	Australia, Oceania	• Decline in soil pH • Weathering enhancement • Early successional peak in nodulation	Chapter 2
<i>Vachellia farnesiana</i>	Leguminosae	tree	yes	Savanna woodland	USA, N America	• Increase in ecosystem N stocks	³⁸
<i>Mimosa arenosa</i>	Leguminosae	tree	yes	Dry tropical forest	Mexico, Mesoamerica		³⁹
<i>Robinia pseudoacacia</i>	Leguminosae	tree	yes	Deciduous temperate forest	USA, N America	• Increase in ecosystem N stocks • Early successional peak in nodulation	⁴⁰
<i>Inga chocoensis</i>	Leguminosae	tree	yes	Wet Tropical forest	Costa Rica, Mesoamerica		⁴¹
<i>Adenanthera pavonina</i> , <i>Albizia lucidior</i> , <i>Albizia vialeneae</i> , <i>Milletia diptera</i> , <i>Azelia xylocarpa</i>	Leguminosae	trees	yes	Wet Tropical forest	Vietnam, Asia		⁴²
<i>Baphia nitida</i>	Leguminosae	tree	yes	Wet Tropical forest	Nigeria, Africa	• Suggested increase in ecosystem N stocks	^{43,44}
<i>Alnus rubra</i>	Betulaceae	tree	yes	Coniferous temperate forest	USA, N America	• Increase in ecosystem N stocks	⁴⁵
<i>Ceanothus velutinus</i>	Rhamnaceae	shrub	yes	Coniferous temperate forest	USA, N America	• Increase in ecosystem N stocks • Early successional peak in nodulation	⁴⁶

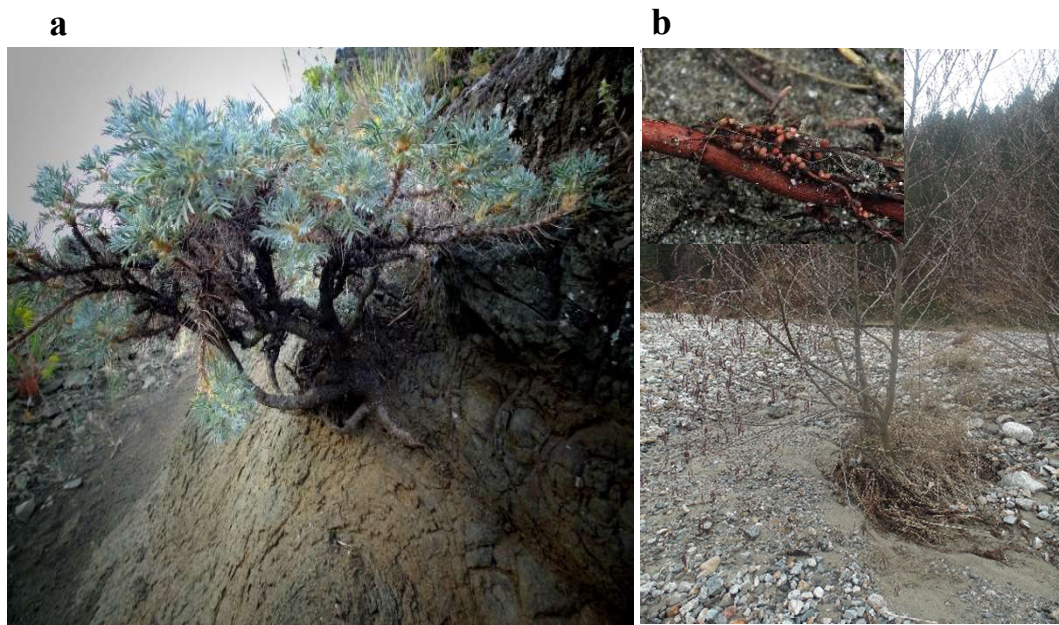


Figure 1. Early pioneering N₂-fixers, such as (a) the legume shrub *Astracantha aitensis* endemic to Bulgaria and (b) the betulaceous *Alnus* sp. (and its *Frankia*-inhabited nodules) illustrated here, are important for soil formation by their N₂-fixing symbioses and active rock-root-microbe interactions. Photo credit: Dimitar Z. Epihov.



Figure 2. Secondary successional N₂-fixers, such as (a) *Inga cocleensis* in Panamanian tropical forests fix large amounts of atmospheric dinitrogen through their (b,c) indeterminate *Burkholderia/Paraburkholderia* and *Bradyrhizobium*-inhabited nodules (see Appendix 4) and (b) actively interact with soil minerals as seen by their highly nodulated roots surrounding a buried bag of soil minerals. Magnification in c is x10. Photo credit: Dimitar Z. Epihov.

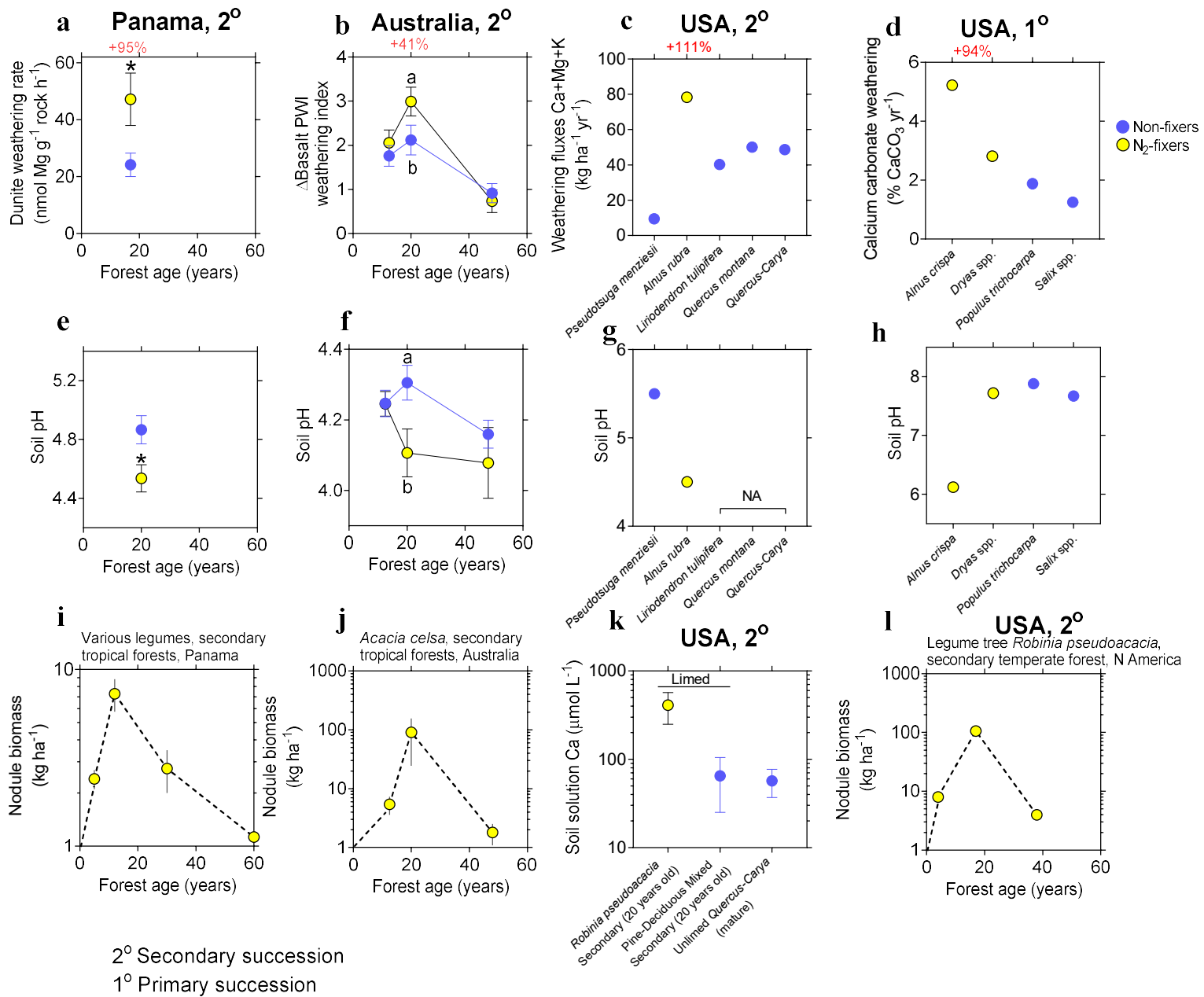
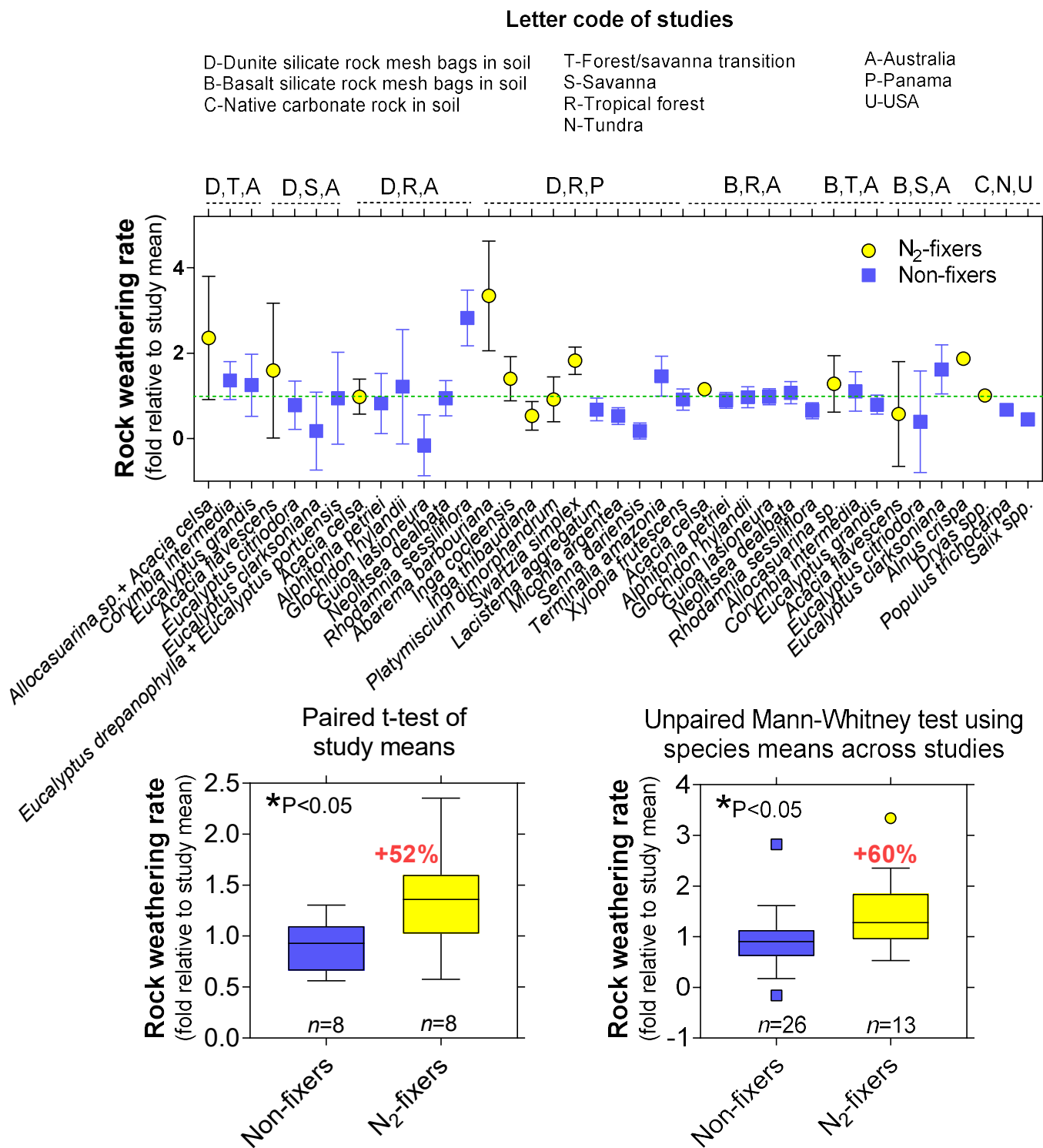


Figure 3. Case studies revealing greater weathering rates and cation export beneath N₂-fixing trees (a-d, k) linked to soil acidification (e-h) and successional-timed nodulation in legume trees (i,j,l). References are as follows: Panama and Australia – Chapters 1 and 2, respectively, USA 2° c,g – Homann et al.¹⁴ and references therein, USA 1° d,h – Crocker and Major³², USA 2° lime/no-lime watershed data from Johnson et al.¹⁵ and references therein, USA 2° nodulation data for *Robinia pseudoacacia* from Boring et al.⁴⁰. Nodulation biomass data for Panamanian legume trees are after Batterman et al.³⁷. Error bars indicate S.E.M. Statistical tests are Mann-Whitney test (P<0.05) and Two-Way ANOVA (P<0.001).



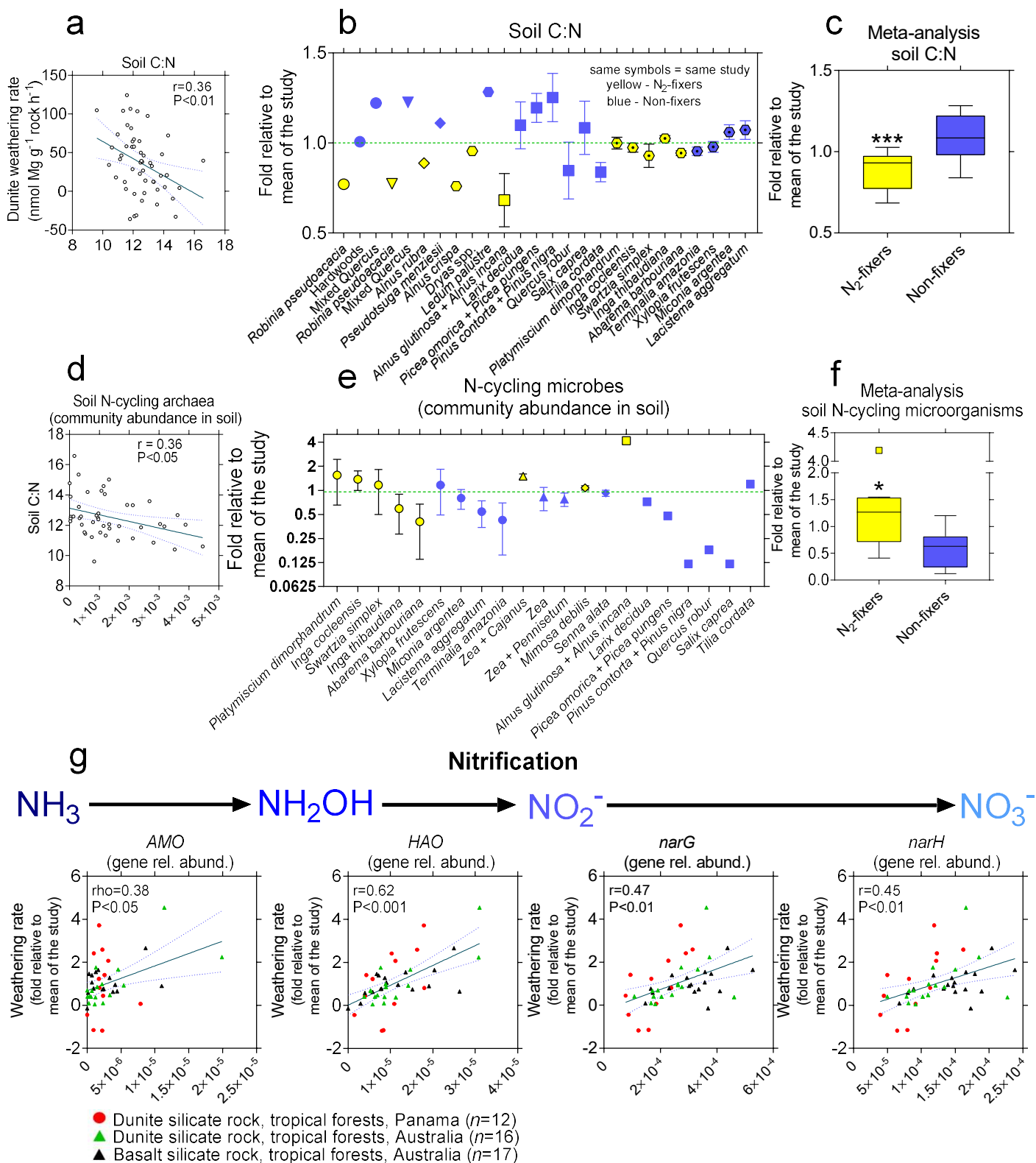


Figure 5. Meta-analyses of biogeochemical and metagenomic data show that the lower soil C:N ratios and higher abundance of N-cycling microorganisms, all indicative of high N inputs by fixers, are linked to weathering rates with strong positive correlations of all major nitrification genes with silicate rock weathering. a. Soil C:N ratios significantly correlate with weathering of dunite in Panamanian tropical forests; **b.** summary of six studies comparing soil C:N in N₂-fixing trees with those beneath non-fixing trees; **c.** meta-analysis using the mean for each species shown in Figure 4b reveals significantly lower C:N in soil of N₂-fixing trees than non-fixers (two-tailed t-test, $P<0.001$); Studies are: circles: USA Boring and Swank, 1984,⁴⁰, triangles: USA Rice et al., 2004,¹⁶, diamonds: USA Cole et al., 1991,⁴⁷,

polygons: USA Rhoades et al., 2001,⁴⁸, squares: Czech Republic Urbanová et al., 2015,⁴⁹, polygons with dot: Panama Epihov et al., 2018, unpublished; **d.** soil N-cycling archaea (combined abundance of the genera *Nitrososphaera* and *Nitrosocaldus*) correlate negatively with soil C:N (Pearson correlation test, $P < 0.05$, $r = -0.36$, $n = 45$) in Panamanian tropical forests; **e.** summary of four studies (circles: Panama – Epihov et al., 2018, unpublished, triangles: Ghana – Sul et al., 2013,⁵⁰, diamonds: Amazonia soil incubation experiment – Barbosa Lima et al., 2015⁵¹, squares: Czech Republic – Urbanová et al., 2015⁴⁹) comparing the abundance of N-cycling microorganisms in N₂-fixing and non-fixing trees (the abundance is the sum of all N-cycling genera including *Nitrososphaera*, *Nitrosocaldus*, *Nitrosomonas*, *Nitrospira*, *Nitrosospira*, *Nitrobacter*, *Denitrobacter*, *Nitratiruptor*, *Nitrosococcus*, *Nitrosovibrio*, *Nitrococcus*, and *Denitromonas* – among which *Nitrospira* and *Nitrososphaera* are dominant accounting for ~59% and 35% of all N-cycling bacteria, respectively, in the Panamanian study – labelled with circular symbols). In the case of the studies labelled with diamonds and squares – only data for the dominant *Nitrospira* was available and used; **f.** meta-analysis of microbial community data using the mean from each species or combination of species in Figure 4e shows that N₂-fixers exhibit significantly greater abundance of N-cycling microorganisms in soil than non-fixing trees (Mann-Whitney test, $P < 0.05$); **g.** all major nitrification genes as identified in KEGG (including *AMO*, *HAO*, *narG*, and *narH*) exhibit significant positive correlations with silicate weathering rates.

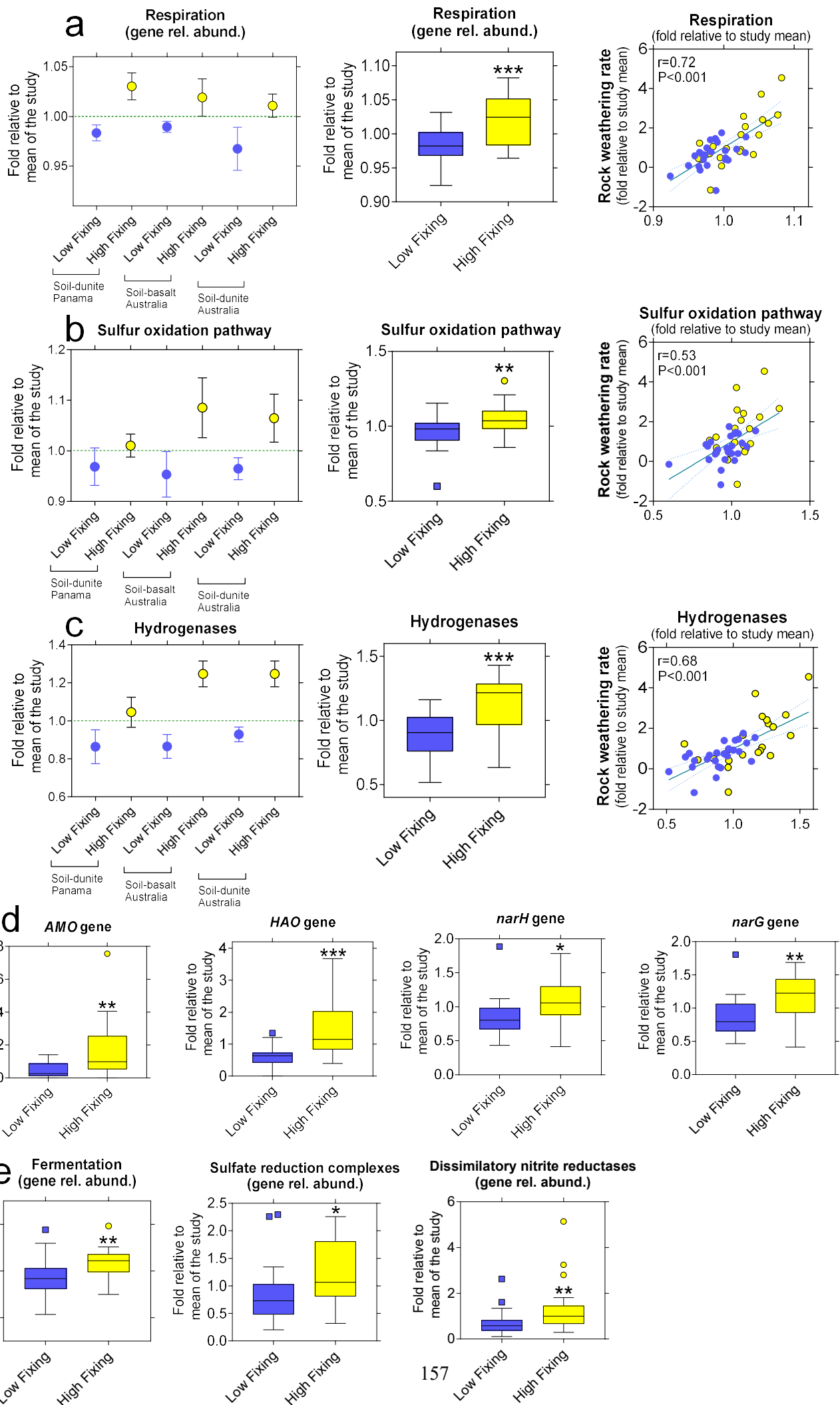


Figure 6. Metagenomic evidence for greater levels of respiratory, lithotrophic (sulfur oxidising, hydrogen oxidising, ammonia/nitrite oxidising) and anaerobic activities in soil-mineral samples from high-fixing forests all correlated to higher rock weathering.

a. The high-level pathway of respiration is enriched in high-fixing relative to low-fixing forests consistent across soil-minerals and studies and correlating with weathering; **b.** The high-level pathway of sulfur oxidation (a type of lithotrophic metabolism) is enriched in high-fixing relative to low-fixing forests consistent across soil-minerals and studies and correlated with rock weathering; **c.** The sum of all hydrogenases genes in the metagenome is greater in high-fixing than low-fixing forests consistent across sites and soil-minerals and correlating with rock weathering rates; **d.** nitrification genes, previously shown to correlate with rock weathering rates (Figure 4g), are significantly more abundant in metagenomes from high-fixing forests relative to low-fixing ones; **e.** metabolic pathways from the anaerobic food chain including fermentation, sulfate reduction and nitrate/nitrite dissimilatory reduction are enriched in high-fixing metagenomes and positively correlate with weathering (not shown). Mann-Whitney test, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$.

Correlations are based on Pearson tests. Error bars in the first figure of each triplet in a, b, c show S.E.M. The whisker box plots contain error bars are based on Tukey distribution.

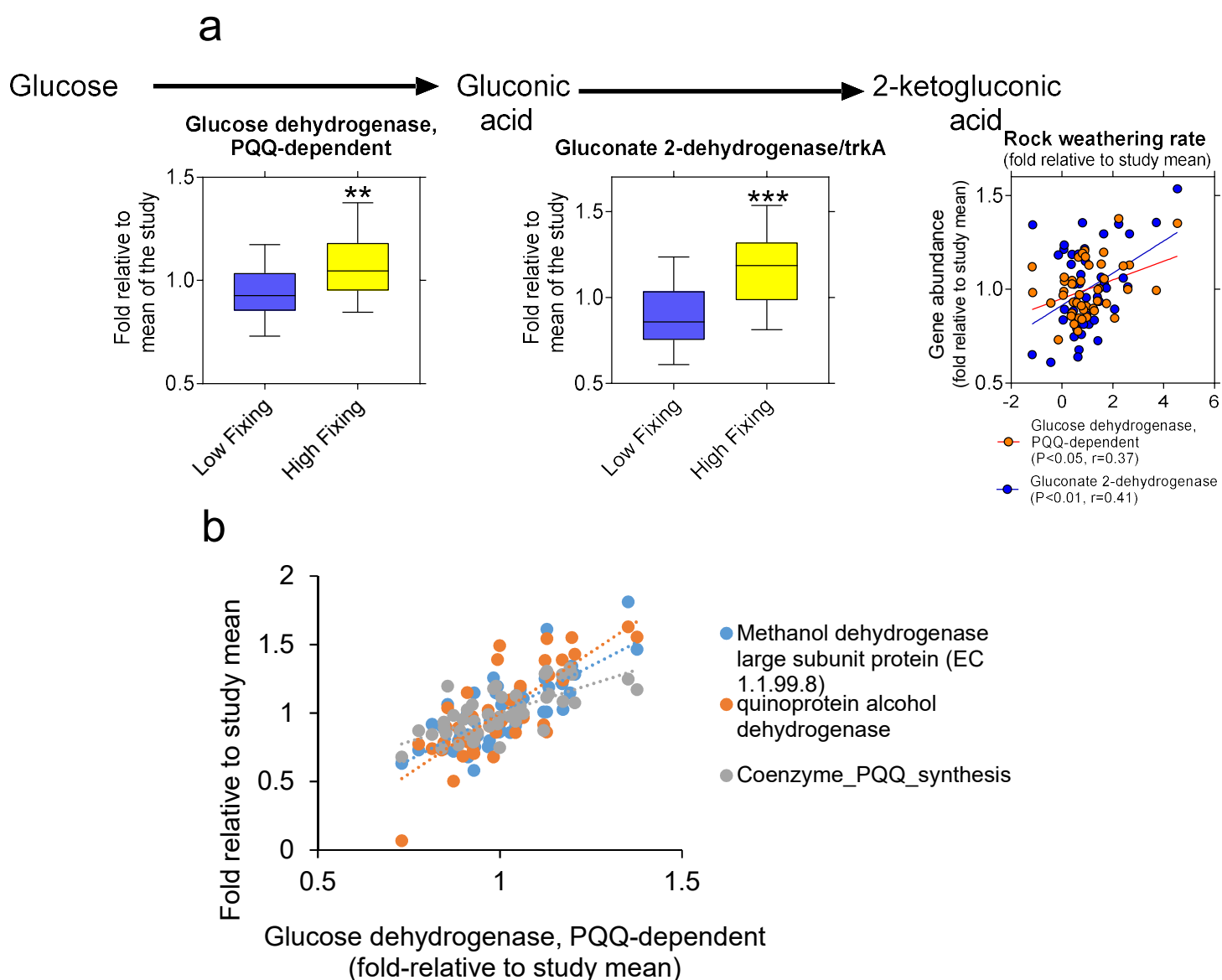


Figure 7. Genes governing the microbial production of gluconic and 2-ketogluconic acid
a. are significantly greater in high-fixing forest samples and correlate with field rock weathering rates and **b.** are dependent on PQQ supply (particularly the *gdhPQQ*, not *trkA*) as evident from co-variation between coenzyme PQQ synthesis pathway and PQQ-dependent enzyme-coding genes (methanol dehydrogenase, quinoprotein alcohol dehydrogenase and *gdhPQQ*) validating the importance of adequate microbial PQQ supply. Statistical tests include Mann-Whitney tests, *** $P < 0.001$, ** $P < 0.01$ and Pearson correlation tests. For **b.** all correlation tests are significant at $P < 0.001$.

ADDED!

Appendices

Appendix 1

Comparing novel and established approaches to taxonomy calling in amplicon libraries of microbial communities

(Annexed to Chapter 1)

Amplicon-based libraries can provide deep structural and functional insights into the microbial communities of soil. Here, we have generated a total of 46 soil rhizospheric libraries containing amplicon DNA sequences of the V4-V5 region of the 16S rRNA gene in prokaryotes. Amplicon libraries were constructed using a two-step PCR approach where the target region was first amplified using locus specific primers with Illumina sequencing primer adapters incorporated into their 5' end (30 cycles), then a second round of PCR was done using forward and reverse primers that contain all Illumina sequencing primer and flow cell binding sequences as well as a unique 8 bp barcode (6 cycles). In the first round of PCR, all samples were amplified in triplicate using the 515F/806R primer set that amplifies the V4-V5 region of the 16S rRNA for bacteria¹. Triplicate reactions were then pooled and unique combinations of barcodes and Illumina adapters added via a second round of PCR. PCR products were purified and normalized using SequalPrep Normalization plates (Life Technologies), pooled into single libraries based on sample type (e.g. soil rhizospheric 16S, soil mineral 16S), concentrated using Agencourt AMPure XP beads, quantified on a Qubit fluorimeter, and quality checked using the High Sensitivity Agilent DNA kit on an Agilent Bioanalyzer. Subsequently, samples were adjusted to appropriate concentrations and sequenced on a total of two runs on an Illumina MiSeq sequencer (2x250 bp paired-ends runs) at the Smithsonian Tropical Research Institute. The generated libraries were de-multiplexed in QIIME² and subsequently processed through 3 separate pipeline processes outlined below:

1. R1 sequences were loaded onto QIIME² closed reference OTU picking using 97% and 90% identity cut-offs and GreenGenes database (<http://greengenes.lbl.gov/Download/>, gg_13_05 release from May, 2013) for the construction of OTU tables (QIIME+GreenGenes+90%identity and QIIME+GreenGenes+97%identity, hereafter **QIG90** and **QIG97**, respectively). Full sequences are aligned (all 250 bp).
2. R1 and R2 sequences were merged and loaded onto the MG-RAST³ server (<http://www.mg-rast.org>) and processed according to default parameters for amplicon libraries. Resulting OTU tables were processed against the in-built SILVA SSU database. As the method of comparison in MG-RAST is blat, the alignment

length is usually much smaller with an average of 103 bp and very high average identity of 98% (hereafter referred to as **MGS**)

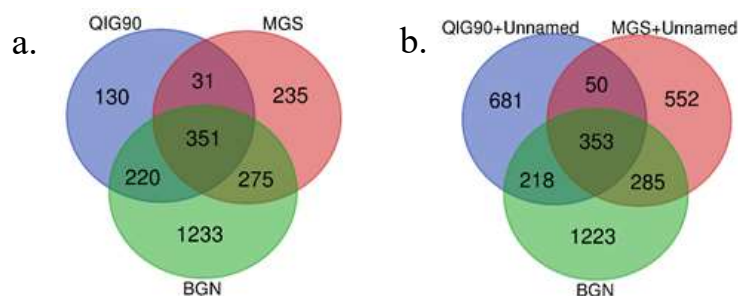
3. R1 and R2 sequences were merged and loaded onto Blast2GO 5 Pro and compared against the NCBI microbial 16S database (accessed in June 2018) using local blastx-fast set at default, except the e-value that was changed from the default e-3 to e-5 and selecting only the single best top ranking hit rather than the top 20 best ranked hits. There was no applied identity cut-off but the average identity was 91% with an average alignment length of ~253bp (hereafter referred to as **BGN**). Higher level of taxonomy of obtained species and genera OTU tables can be obtained through manual assignment in the Taxonomy Summary tool supported by KEGG (<https://www.genome.jp/tools-bin/taxsummary>).

Rarefaction curves revealed satisfactory levels of sequencing depth in all 46 samples and no samples were excluded from further analyses.

Comparisons of the three workflows

Coverage of class and genus level OTUs

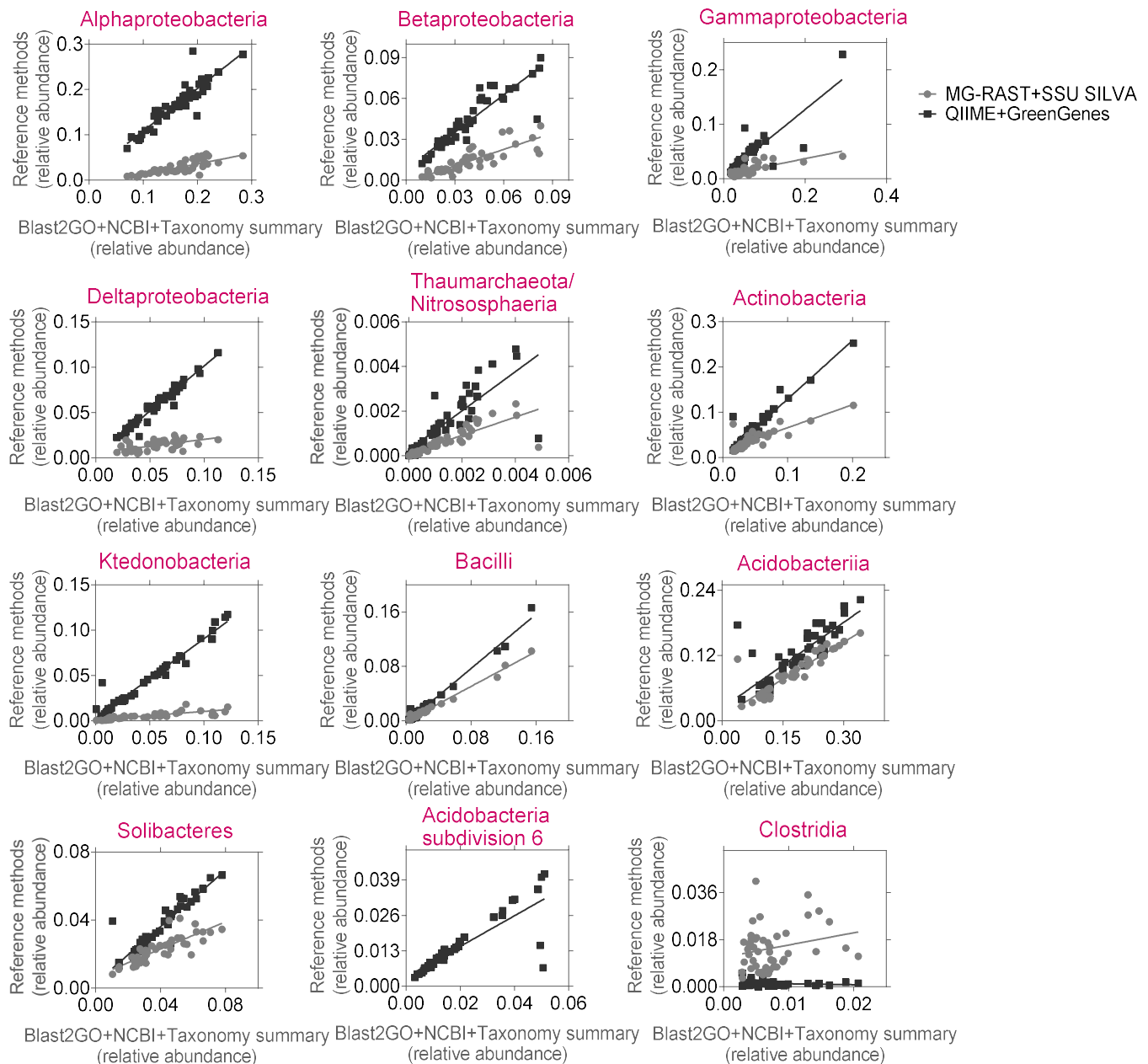
BGN generated the greatest number of named + unnamed (“derived from...”; novel) genera (2079) followed by QIG90 (1302), MGS (1240), and QIG97 (1127). Excluding any unnamed genera derived from family/order/class etc., BGN generated the same number, 2079 genera, as it does not bin genera into derived categories, MGS – 892, QIG90 – 732 and QIG97 – 459 (**Appendix Figure 1**). Such differences in the number of named genera can result from strict cut-offs (as in QIG97) but also from differences in size of the selected databases as QIG90 contained still smaller number of genera than BGN despite the average genus identity in BGN output of ~88% (**Appendix Table 2**). For instance, the NCBI 16S microbial database used in our BGN approach covers ~65-fold and 10-fold more genera than GreenGenes and SILVA SSU, respectively, with authors recommending the further use of the NCBI database in more sequencing microbiome studies⁴.



Appendix Figure 1. Number of unique and overlapping (a) named genera only and (b) named and unnamed (novel; “derived from...”) genera in 46 rhizospheric soil samples analyzed by the three taxonomic calling approaches. QIG90 = QIIME + GreenGenes at 90% identity cut-off, MGS=MG-RAST+SILVA SSU, BGN=BlastGO+NCBI. Note QIG97 not shown in this diagram.

Despite their differing numbers of genera, the relative abundance of 11 out of the 12 analysed major prokaryotic classes significantly correlated with BGN and QIG97 relating better than the abundance generated by MGS (**Appendix Figure 2**). The average identity of genera in BGN is 91% which may seem too relaxed relative to the recommended 95% genus

Class-level comparisons



Appendix Figure 2. Class-level comparisons reveal that the newly proposed BGN method (Blast2GO+NCBI+Taxonomy summary) agrees well the output of well-established analogues QIIME+GreenGenes at 97% identity cut-off (QIG97) and MG-RAST+SILVA SSU (MGS).

cut-off⁵ suggesting that any hits below 95% may represent (case 1) a novel closely related genus or perhaps (case 2) be part of the same genus if the latter is with high-intragenic variation. We find that the relative abundance of over 47% of all genera shared between QIG97 and BGN correlate significantly. More than half of those reveal average identity <95%. Combined these findings suggest that either these are part of the same genus (OTU_{X1} is a species of genus X). An alternative explanation is that OTU_{X1} and X are separate genera that share similar physiology or niches in soil causing their apparent correlation at even lower identity thresholds (OTU_{X1} is part of a novel genus *X1* that is most closely related to genus X). Although OTU_{X1} and genus X may differ in certain aspects of their physiology, functionally in the context of the soil community, OTU_{X1} and genus X can be seen as “behaving” as a single unit supported both by their correlation and sequence based similarity.

Methods such as QIG bin OTUs that are below a certain cut-off in “derived from family/order/class etc.” so that OTU_{X1}, OTU_{X2}...OTU_{Xn} revealing 93%<identity<95% would all be part of a “derived from family” cluster. Consequently, no functional information is known about these as they will represent a collection of sequentially distinct units as demonstrated in the following example below:

Family M = {genus U, genus V, genus W, genus X, genus Y, genus Z}

OTU_{X1} best hit: 94.1% identity with genus X → binned in “derived from family M”

OTU_{X2} best hit: 94.0% identity with genus X → binned in “derived from family M”

OTU_{X3} best hit: 93.7% identity with genus Y → binned in “derived from family M”

...

OTU_{X7} best hit: 93.6% identity with genus Z → binned in “derived from family M”

Although OTU_{X1} and OTU_{X2} are most closely related to each other and to X than to any of the other genera in family M, they are all binned together with OTU_{X3} and OTU_{X7} into “derived from family M” making their functional differentiation impossible as the resulting “derived from family M” would carry a mixture of X-like, Y-like and Z-like traits.

However, if a database is relatively limited in their number of genera and family M is not fully represented by all of its constituent genera so that:

Family M = {genus U, genus Y, genus Z}

and

OTU_{X1} external BLAST best hit: 94.1% identity with genus X; second best hit 89.2% identity with genus Y → binned in “derived from family M”

OTU_{X2} external BLAST best hit: 94.0% identity with genus X; second best hit 88.4% identity with genus Y → binned in “derived from family M”

binning OTUs in a “derived from...” would be more appropriate than binning them in genus Y.

However, databases that cover more of the currently supported phylogeny (such as NCBI, which is the largest sequence-containing microbial taxonomy database⁴) would be less constrained and therefore should be able to place OTU_{X1-X7} in their respective most closely related genera rather than clustering them all in one or two more distantly related genera. Indeed, upon discovery and description of all bacterial genera, local blasts against such a future complete database would yield exact matches thus eliminating the need of “derived from...” clusters regardless of % identity.

Even at present, while such a database is yet to materialise, substantially large databases such as NCBI in combination with no imposed identity cut-offs should be able to provide an enhanced insight into the functional traits of the analysed 16S-based microbiomes while lacking information on potentially new taxa. In contrast, methods relying on strictly embedded identity cut-offs (98.65% for species and 95% for genus^{5,6}) will be bound to present an exaggerated estimate of novel taxa. For example, some such predictions yield over 3600 new genera in a single soil community sample⁷. However, analysis of over 150 genera reveals that 62% of these contain published species that exhibit lower than the accepted 95% identity with other members of the same genus. Different monophyletic clades vary substantially in their level of 16S rRNA gene conservation and taxonomy based on strict thresholds equal for all taxa may consequently be misleading⁵. As a result, we hypothesize that the BGN approach, proposed here, may be better suited to functional characterisation of environmental microbial communities.

Comparison of BGN vs. QIG vs. MGS in their ability to predict function in the microbiome

Linking microbial abundance to functional potential is challenging but has been previously attempted. One such example is the platform PICRUSt⁸ which can reconstruct abundance profiles for KEGG gene orthologues using 16S rRNA closed-reference OTU picking in QIIME and its resulting BIOM tables. However, one limitation of that approach is the large gap between number of genera in 16S databases and the number of sequenced genomes. Similarly, BLAST-ing annotated metagenomic sequences to establish the taxonomic identity of genera contributing to the pool of a particular gene of interest will also be hampered by the comparatively small number of sequenced genomes as well as by the often highly conserved sequences of such genes.

Alternatively, another approach is to utilise detailed literature search in generating large lists of genera that have been documented to perform a particular function of interest – e.g. nitrification, sulfur reduction, iron reduction, methanogenesis etc. Consequently, manual extraction of genera from OTU tables and summing of their abundance should serve as a

proxy proportional to the abundance of marker genes in the metagenome that are involved in the same function or process.

Because such a methodology would ultimately rely on reports of a given phenotypic trait in a genus, OTU tables containing large number of genera such as BGN may be better equipped in linking structure to function. In support of that statement, we find that the summed abundance of genera in BGN could significantly link the abundance of ammonia oxidisers to ammonia monooxygenase (*AMO*) gene abundance ($n = 12$ 16S rRNA libraries paired with $n = 12$ shotgun metagenomic libraries), nitrite oxidisers to hydroxylamine reductase (*HAO*), Fe(III) reducing genera to Fe(III) respiration - *Shewanella* type

Appendix Table 1. Inferred links between structure of the microbiome (relative abundance of taxa) and function in the metagenome (relative abundance of marker genes or pathways of genes).

Sum of 16S rRNA relative abundance	BGN	QIG97	MGS	Metagenomic gene markers
Ammonia oxidising genera	**	*	ns	<i>AMO</i>
Nitrite oxidising genera	***	***	*	<i>HAO</i>
Fe(III) reducing genera	**	ns	•	Fe(III) respiration - <i>Shewanella</i> type (all genes)
Sulfur cycling genera	* _{Sp}	ns	ns	Inorganic sulfur assimilation (all genes)
Methanogenic genera	ns	ns	ns	<i>PMO/MMO</i>

AMO – ammonia monooxygenase; *HAO* – hydroxylamine reductase; *PMO+MMO* – particulate methane monooxygenase and methane monooxygenase. Pearson test, *** $P < 0.001$, ** $P < 0.01$, * $P < 0.05$, • $P < 0.10$, ^{ns} $P > 0.10$, except where the test was Spearman due to lack of normal distribution (denoted with “_{Sp}”).

pathway, and sulfur cycling genera to inorganic sulfur assimilation (**Appendix Table 1**). In comparison, in QIG97 correlation was only observed between genera abundance with *AMO* and *HAO*; and MGS-generated abundance profiles only correlated with *HAO* and Fe(III)

respiration genes (the latter correlated at Pearson test $P < 0.10$). Neither one of the three methods was able to link their assigned methanogen abundance to the abundance of methane monooxygenase genes.

These findings suggest that taxonomy calling by BGN generates an enhanced approximation of soil microbiomes in relation to important functional traits. However, certain processes still remain elusive to such approaches (e.g. methanogenesis) indicating that further research in the lineages involved is required before taxonomy can reliably be translated into function.

Appendix Table 2. Top 60 Blast2GO genus-level OTUs – their average e-value, mean alignment length (base pairs), total number of hits and mean identity % as well as the same across all genera and hits (at the bottom of the table) in soil microbiome samples 32-41 ($n = 10$).

Genus	Mean E-Value	Mean alignment length (bp)	Total number of hits	Mean identity %
<i>Chthoniobacter</i>	1.54E-44	256	39901	91.5
<i>Acidobacterium</i>	5.48E-57	250	25584	91.1
<i>Limisphaera</i>	6.07E-18	260	22606	89.0
<i>Edaphobacter</i>	5.22E-48	250	13187	88.3
<i>Ktedonobacter</i>	6.83E-55	253	12037	85.1
<i>Occallatibacter</i>	2.33E-65	252	10926	89.4
<i>Bradyrhizobium</i>	2.9E-80	256	10778	97.8
<i>Paludibaculum</i>	1.41E-49	260	10440	89.5
<i>Vicinamibacter</i>	1.06E-63	265	10079	91.6
<i>Rhodoplanes</i>	4.04E-82	257	9116	96.9
<i>Gemmata</i>	6.49E-58	259	8979	87.4
<i>Burkholderia</i>	5.57E-63	262	6609	98.3
<i>Bacillus</i>	3.09E-22	254	5338	97.7
<i>Acidibacter</i>	2.98E-59	257	4889	95.5
<i>Pseudolabrys</i>	7.08E-57	262	4872	96.5
<i>Pseudomonas</i>	3.37E-65	255	4217	98.1
<i>Rhizomicrobium</i>	4.6E-73	268	3860	93.1

<i>Granulicella</i>	6.78E-16	252	3824	88.6
<i>Actinoallomurus</i>	2.23E-69	253	3277	95.2
<i>Desulfonatronum</i>	9.41E-45	254	3213	87.0
<i>Streptomyces</i>	6.32E-71	258	3146	97.3
<i>Chryseolinea</i>	1.03E-24	263	3069	91.1
<i>Pseudorhodoplanes</i>	8.64E-71	259	3063	96.6
<i>Bryobacter</i>	3.2E-69	264	3034	89.8
<i>Skermanella</i>	4.59E-68	272	2988	91.2
<i>Zavarzinella</i>	6.66E-60	269	2863	87.5
<i>Gaiella</i>	3.32E-67	262	2734	92.3
<i>Azospirillum</i>	7.52E-28	258	2723	92.1
<i>Stenotrophobacter</i>	1.4E-62	287	2557	94.3
<i>Thauera</i>	3.41E-64	280	2536	92.5
<i>Tepidisphaera</i>	4.37E-51	272	2472	87.8
<i>Sphingomonas</i>	2.04E-77	268	2461	97.8
Rhizobium	3.97E-20	258	2437	98.0
<i>Sphingobium</i>	4.95E-85	255	2357	96.1
<i>Stella</i>	4.41E-25	256	2342	92.0
<i>Klebsiella</i>	2E-99	253	2273	99.3
<i>Silvibacterium</i>	7.05E-61	252	2219	94.0
<i>Gemmatimonas</i>	5.86E-62	280	2195	88.9
<i>Thermostilla</i>	1.33E-72	280	2136	89.6
<i>Fimbriiglobus</i>	1.56E-63	268	2111	88.5
<i>Terrimonas</i>	5.19E-79	267	2093	96.1
<i>Mucilaginibacter</i>	1.13E-08	256	2090	98.2
<i>Terrimicrobium</i>	8.88E-54	256	1904	89.0
<i>Nevskia</i>	1.13E-59	290	1687	92.9
<i>Niastella</i>	3.26E-79	257	1673	95.5

<i>Haliangium</i>	3.97E-69	263	1644	90.5
<i>Parasediminibacterium</i>	1.55E-64	257	1595	94.6
<i>Phenylobacterium</i>	4.7E-65	255	1513	96.3
<i>Aliidongia</i>	3.09E-77	271	1507	90.6
<i>Planctopirus</i>	1.3E-59	279	1462	88.8
<i>Reyranella</i>	2.19E-26	266	1432	96.2
<i>Paraburkholderia</i>	1.19E-66	257	1408	97.8
<i>Niabella</i>	7.66E-77	275	1318	95.0
<i>Geobacter</i>	1.28E-66	259	1308	91.0
<i>Massilia</i>	2.37E-63	257	1279	92.8
<i>Anaeromyxobacter</i>	3.75E-67	257	1279	90.8
<i>Kofteria</i>	4.03E-52	260	1265	90.1
<i>Brevitalea</i>	4E-30	260	1265	92.6
<i>Labilithrix</i>	5.9E-69	261	1264	92.6
<i>Fimbriimonas</i>	1.77E-69	284	1226	88.4
...
	<i>Median</i>	<i>Average</i>		<i>Average</i>
Across genera	1.01E-70	250		87.8
Across all hits	1.7E-101	259		91.6

Note that all legume nodulating genera found in top 60 are highlighted in green, with all of them exhibiting identity >95% consistent with the accepted genus-level cut-off⁵.

Appendix 2

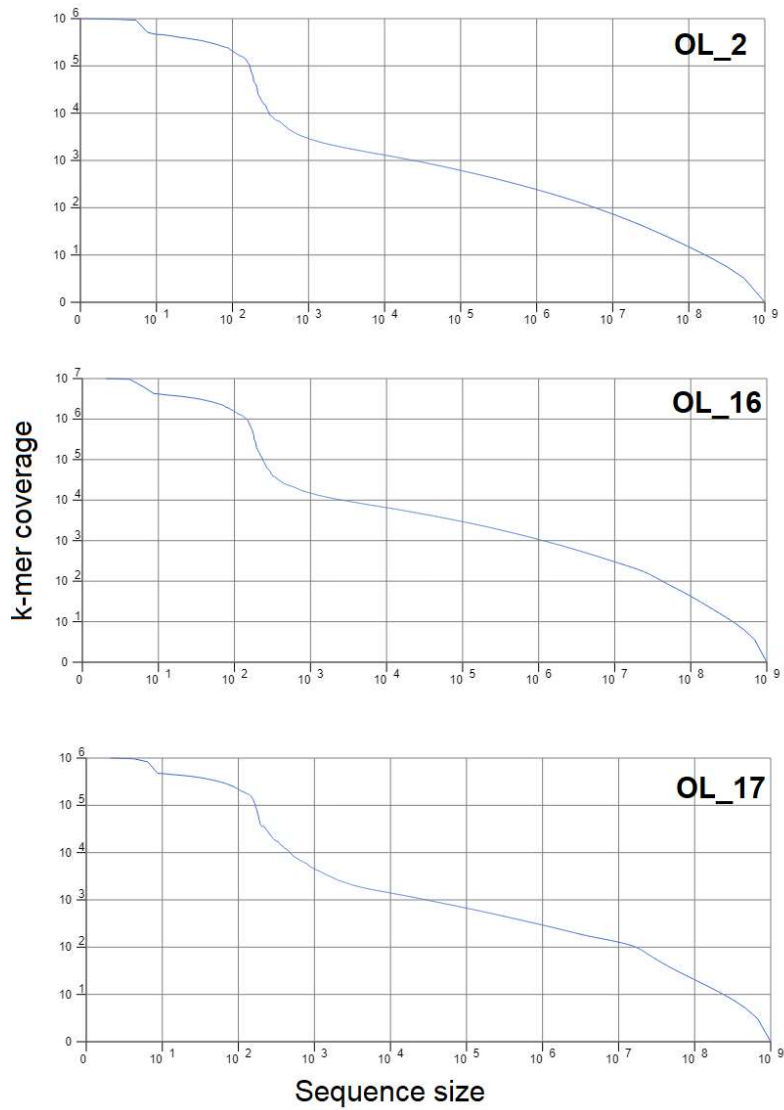
Shotgun Metagenome Statistics

(Annexed to Chapters 1 and 2)

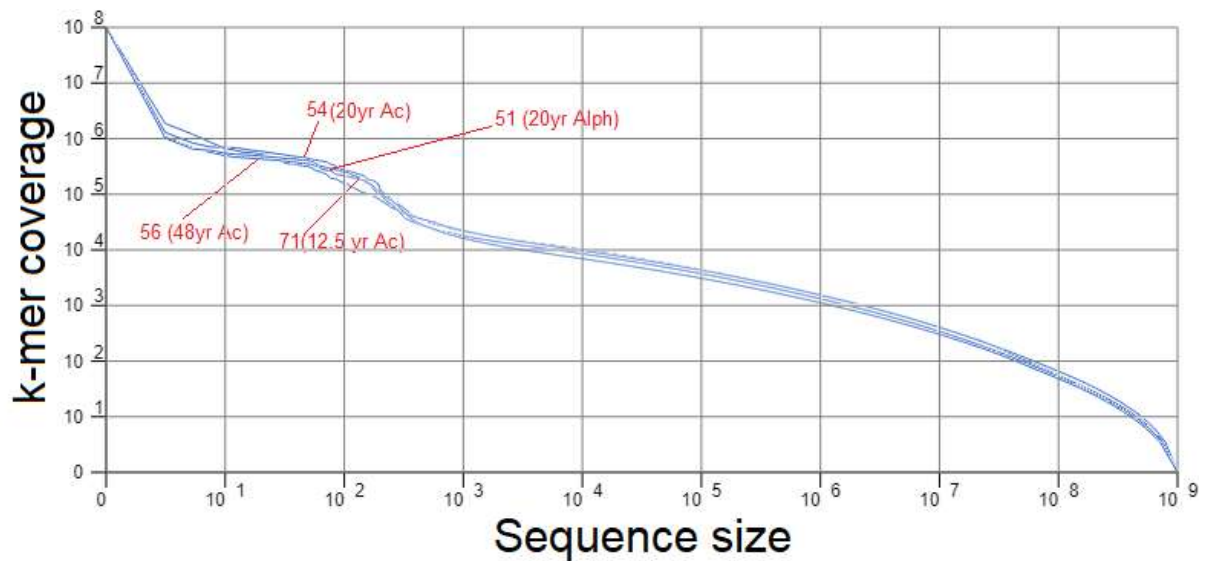
Appendix Table 3. Shotgun metagenome libraries from our tropical forest weathering study in Panama (Chapter 1) and their MG-RAST ID numbers and size

MG-RAST ID	Name	Base pairs (bp) count	No. of sequences
mgm4750411.3	OL_2	2,868,154,772	17,959,686
mgm4750538.3	OL_1	6,067,791,629	38,077,834
mgm4750539.3	OL_7	4,935,659,002	30,420,129
mgm4750544.3	OL_16	6,825,659,192	42,207,195
mgm4751389.3	OL_17	5,642,423,815	35,154,753
mgm4751392.3	OL_31	4,626,468,473	28,844,519
mgm4751395.3	OL_29	4,501,728,547	28,118,622
mgm4751396.3	OL_33	2,891,291,381	18,057,850
mgm4751576.3	OL_47	3,275,979,911	20,550,000
mgm4751577.3	OL_48_	4,413,291,146	27,590,400
mgm4751584.3	OL_50	4,075,326,761	25,439,339
mgm4751599.3	OL_58	7,132,861,445	44,185,203
mgm4751712.3	S_17_	7,793,811,849	48,374,646
mgm4751755.3	S_29	7,823,595,015	48,433,318
mgm4752317.3	S_33	7,942,191,469	49,059,560

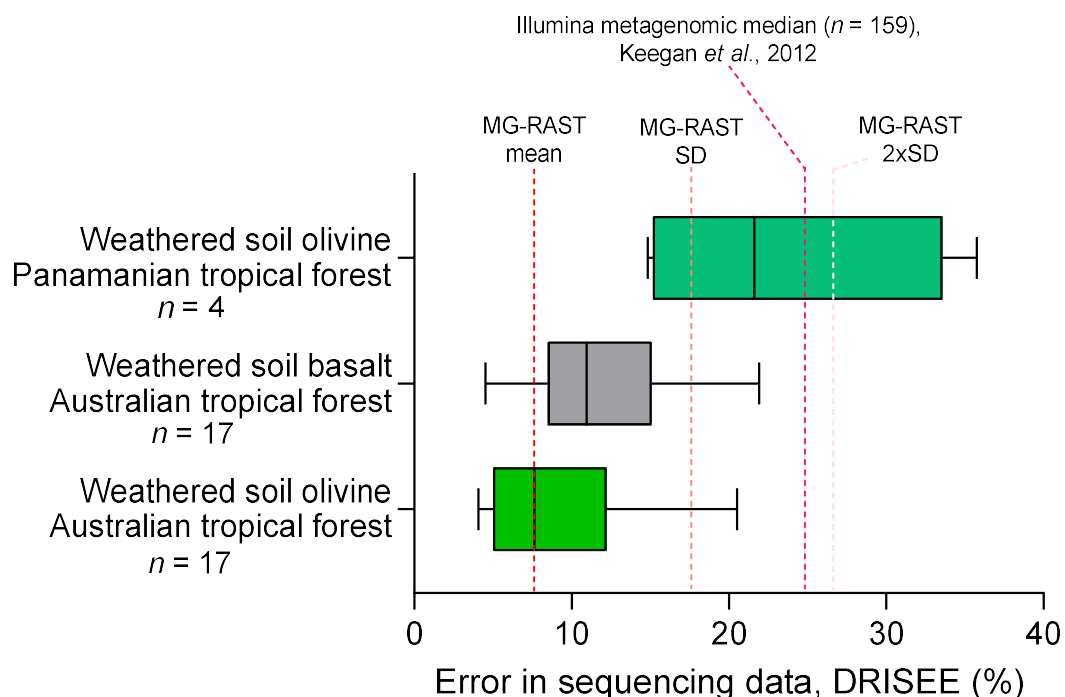
OL – weathered soil olivine samples; S – soil rhizosphere samples. Note that the number of sequences of weathered soil minerals from beneath N₂-fixers (OL17-OL48), non-fixers near N₂-fixers (NF-near: OL1, 2, 50) and non-fixers far from N₂-fixers (NF-far: OL7, 16, 58) do not differ significantly neither in number of sequence reads (Welch's ANOVA, P>0.10), nor in total number of base pairs (Welch's ANOVA, P>0.10).



Appendix Figure 3. k-mer sequencing coverage is similar across shotgun metagenomes of weathered soil mineral from beneath different functional groups of trees in Panamanian tropical forest (Chapter 1) indicating comparable sequencing depths.



Appendix Figure 4. k-mer sequencing coverage is similar across shotgun metagenomes of weathered soil olivine mineral from beneath N₂-fixing *Acacia celsa* (Ac) and non-fixing *Alphitonia petriei* (Alph) trees at different ages of secondary forest succession (Chapter 2) indicating comparable sequencing depths.



Appendix Figure 5. DRISEE sequencing error distribution in each of the three groups of shotgun metagenomes (Australian tropical forest soil basalt and soil olivine – Chapter 2 and Panamanian tropical forest soil olivine – Chapter 1) shows that our libraries are largely similar to other metagenomes hosted at the MG-RAST server (with means falling within the range defined by the mean and 2 x standard error) and those compiled by the original DRISEE paper⁹ by Keegan *et al.*, 2012. The Australian shotgun metagenomes (prepared using the TruSeq Nano gel free kit and sequenced on an Illumina NovaSeq) reveal lower DRISEE errors than the Panamanian shotgun metagenomes (prepared using Nextera

XT library prep-up kit and sequenced on an Illumina HiSeq machine). Note that for shotgun metagenomes of weathered soil olivine samples in Panama $n = 4$ because MG-RAST failed to generate DRISSE values for the remaining 8 samples. Note that analyses of DRISSE values between tree groups (non-fixers and N_2 -fixers) and between different forest ages (12.5, 20, 48 years old forests) yielded no significant differences between these groups neither in olivine, nor in basalt from the soil of Australian tropical forests in Chapter 2 (two-tailed t-test, $P > 0.10$; Welch's ANOVA, $P > 0.10$).

Appendix 3

Amplicon library sequencing statistics

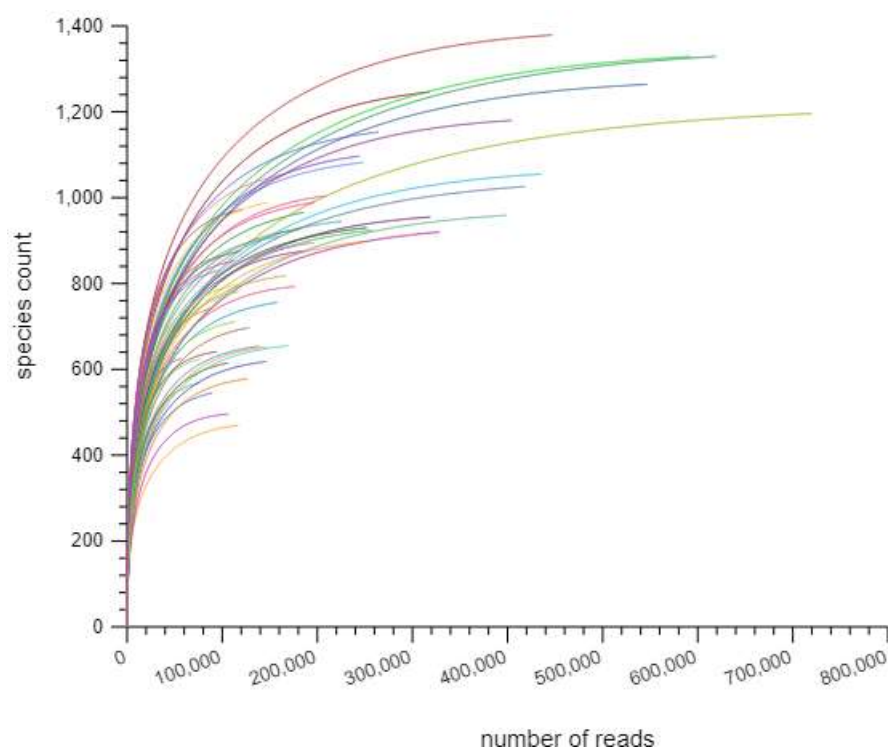
(Annexed to Chapter 1)

Appendix Table 4. The 46 soil rhizospheric samples and their sequencing statistics

MG-RAST ID	Sample name	Base pairs (bp) count	Number of sequences
mgm4756263.3	1	63,854,634	252,483
mgm4756256.3	2	40,022,945	158,114
mgm4756275.3	3	42,858,015	169,309
mgm4756277.3	4	42,312,921	167,186
mgm4756267.3	5	49,770,373	196,653
mgm4756266.3	6	44,649,079	176,577
mgm4756274.3	7	40,438,161	159,796
mgm4756284.3	8	41,769,319	165,041
mgm4756254.3	9	62,893,261	248,475
mgm4756264.3	10	30,252,288	119,392
mgm4756246.3	11	46,130,351	182,257
mgm4756278.3	12	32,551,375	128,519
mgm4756258.3	13	57,141,266	225,702
mgm4756286.3	14	52,771,514	208,488
mgm4756242.3	15	66,947,090	264,442

mgm4756271.3	16	47,073,898	186,072
mgm4756262.3	17	49,710,367	196,502
mgm4756281.3	18	36,896,361	145,811
mgm4756250.3	19	65,169,337	257,393
mgm4756251.3	20	63,143,293	249,501
mgm4756270.3	24	30,645,778	121,093
mgm4756260.3	25	28,577,246	112,909
mgm4756243.3	26	36,937,032	145,587
mgm4757411.3	27	29,671,543	117,235
mgm4756269.3	29	61,957,368	244,705
mgm4756247.3	30	26,976,030	106,574
mgm4756261.3	31	26,973,619	106,566
mgm4757419.3	33	37,579,765	148,565
mgm4757416.3	34	19,565,480	77,200
mgm4757417.3	36	35,092,277	138,680
mgm4756245.3	37	22,687,404	89,666
mgm4757415.3	39	23,914,482	94,429
mgm4757420.3	40	25,700,294	101,522
mgm4756252.3	41	32,083,674	126,763
mgm4756249.3	43	14,451,495	57,087
mgm4757421.3	45	29,658,137	117,153
mgm4756265.3	46	28,802,145	113,782
mgm4756280.3	47	105,979,767	418,805
mgm4756285.3	48	150,092,873	593,223
mgm4757412.3	50	101,003,730	398,844
mgm4757422.3	51	83,243,133	328,719
mgm4757418.3	52	182,368,519	720,242
mgm4756257.3	53	80,479,270	317,881

mgm4756283.3	54	102,373,708	404,467
mgm4757414.3	56	110,218,169	435,625
mgm4756287.3	57	138,345,533	546,603



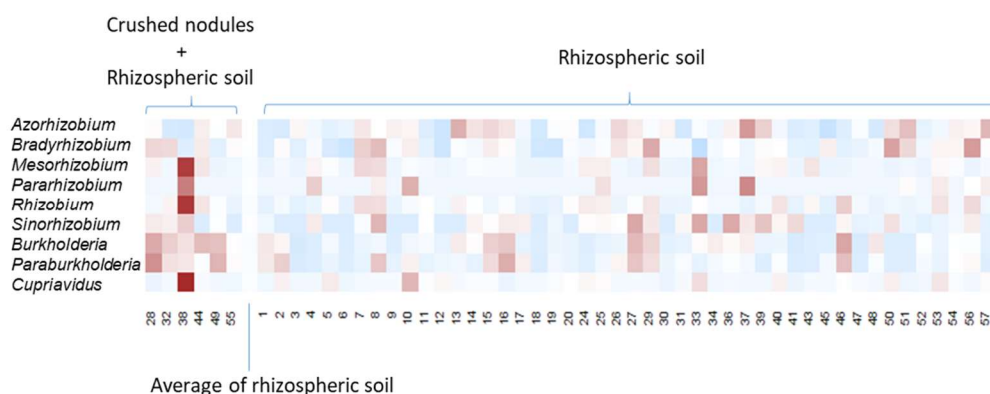
Appendix Figure 6. Rarefaction curve generated in MG-RAST using the SILVA SSU database for the soil microbial community samples demonstrates sufficient sequencing depth with the majority of amplicon libraries near to or reaching a plateau.

Appendix 4

Symbiotic bacteria in crushed nodule and rhizosphere mixed samples

(Annexed to Chapter 1 and Chapter 4)

Our secondary tropical forest sites represent a diverse mixture of nodulating legumes, but *Inga* trees are the most predominant in terms of numbers and basal tree area, particularly early-pioneering *Inga cocleensis* (sample 28; **Appendix Figure 7**) and *Inga thibaudiensis* (sample 32; samples 38 – *Platymiscium dimorphandrum*, 44 – *Swartzia simplex*, 49 and 55 – nodules from the rhizosphere of non-legumes with unknown legume hosts; **Appendix Figure 7**). To gain insight into the nodulating bacteria of legumes in our sites, we sequenced and performed taxonomy calling of 16S rRNA amplicon libraries of soil rhizosphere + crushed nodules ($n = 6$) and compared them against samples only containing soil rhizosphere communities ($n = 46$). Our analyses reveal a major >3-fold enrichment in the atypical β -proteobacterial nodulating genera *Burkholderia* and *Paraburkholderia* as well as a more minor 1.5-fold enrichment of *Bradyrhizobium* in *Inga* nodules (samples 28 and 32) relative to soil only samples (**Appendix Figure 7**). While the majority of nodulating reports from *Inga* trees indicate symbiotic interactions mainly with *Bradyrhizobium*^{10,11,12}, a study has also extracted several *Burkholderia* isolates from nodules of *Inga vera* in the Atlantic rainforest in Brazil¹³. Therefore, our preliminary findings are consistent with the published literature but highlight the need of further research into non-bradyrhizobial nodulating agent within this diverse and dominant Neotropical genus of legumes.



Appendix Figure 7. Heatmap of the relative abundance of different nodulating genera in crushed nodules+rhizospheric soil versus rhizospheric soil only 16S rRNA libraries. Note that the colour coding is as follows: high abundance – dark red, medium abundance – white, low abundance – blue).

Little information exist on nodulation in *Platymiscium* with previous research indicating extraction of very slow to slow growing isolates from nodules without characterising their taxonomic identity¹⁴. Here, we report that in addition to enrichment in *Burkholderia* and *Paraburkholderia*, as well as in *Mesorhizobium*, *Rhizobium* and *Sinorhizobium*, nodules from the faboid *P. dimorphandrum* exhibited high abundance of *Cupriavidus*, another atypical rhizobial genus that was not enriched in any other nodule samples (**Appendix Figure 7**). As studying the detailed species-level resolution of nodulating agents was beyond the scope of our study, our generated V4-V5 16S rRNA sequences can only provide reliable differentiation at the genus level (see **Appendix Table 2** and text within **Appendix 1**) but it, nevertheless, reveals that interactions between tropical forest legume trees and rhizobia may not be as simple as “single species-single host scenario” and require further detailed study using other more suitable marker genes as *recA* and *nifH* (currently underway; Brandt, P. and Hedin, L; personal communication).

Appendices references

1. Fierer, N., Leff, J.W., Adams, B.J., Nielsen, U.N., Bates, S.T., Lauber, C.L., Owens, S., Gilbert, J.A., Wall, D.H. and Caporaso, J.G., 2012. Cross-biome metagenomic analyses of soil microbial communities and their functional attributes. *Proceedings of the National Academy of Sciences*, 109(52), pp.21390-21395.
2. Caporaso, J.G., Kuczynski, J., Stombaugh, J., Bittinger, K., Bushman, F.D., Costello, E.K., Fierer, N., Pena, A.G., Goodrich, J.K., Gordon, J.I. and Huttley, G.A., 2010. QIIME allows analysis of high-throughput community sequencing data. *Nature methods*, 7(5), p.335.
3. Meyer, F., Paarmann, D., D'Souza, M., Olson, R., Glass, E.M., Kubal, M., Paczian, T., Rodriguez, A., Stevens, R., Wilke, A. and Wilkening, J., 2008. The metagenomics RAST server—a public resource for the automatic phylogenetic and functional analysis of metagenomes. *BMC bioinformatics*, 9(1), p.386.
4. Balvočiūtė, M. and Huson, D.H., 2017. SILVA, RDP, Greengenes, NCBI and OTT—how do these taxonomies compare?. *BMC genomics*, 18(2), p.114.
5. Rossi-Tamisier, M., Benamar, S., Raoult, D. and Fournier, P.E., 2015. Cautionary tale of using 16S rRNA gene sequence similarity values in identification of human-associated bacterial species. *International journal of systematic and evolutionary microbiology*, 65(6), pp.1929-1934.
6. Kim, M., Oh, H.S., Park, S.C. and Chun, J., 2014. Towards a taxonomic coherence between average nucleotide identity and 16S rRNA gene sequence similarity for species demarcation of prokaryotes. *International journal of systematic and evolutionary microbiology*, 64(2), pp.346-351.
7. Edgar, R.C., 2018. Accuracy of taxonomy prediction for 16S rRNA and fungal ITS sequences. *PeerJ*, 6, p.e4652.
8. Langille, M.G., Zaneveld, J., Caporaso, J.G., McDonald, D., Knights, D., Reyes, J.A., Clemente, J.C., Burkepille, D.E., Thurber, R.L.V., Knight, R. and Beiko, R.G., 2013. Predictive functional profiling of microbial communities using 16S rRNA marker gene sequences. *Nature biotechnology*, 31(9), p.814.
9. Keegan, K.P., Trimble, W.L., Wilkening, J., Wilke, A., Harrison, T., D'Souza, M. and Meyer, F.,

2012. A platform-independent method for detecting errors in metagenomic sequencing data: DRISSE. *PLoS computational biology*, 8(6), p.e1002541.
10. Leblanc, H.A., McGraw, R.L., Nygren, P. and Le Roux, C., 2005. Neotropical legume tree *Inga edulis* forms N₂-fixing symbiosis with fast-growing *Bradyrhizobium* strains. *Plant and soil*, 275(1-2), pp.123-133.
 11. da Silva, K., De Meyer, S., Rouws, L.F., Farias, E.N., dos Santos, M.A., O'Hara, G., Ardley, J.K., Willems, A., Pitard, R.M. and Zilli, J.E., 2014. *Bradyrhizobium ingae* sp. nov., isolated from effective nodules of *Inga laurina* grown in Cerrado soil. *International journal of systematic and evolutionary microbiology*, 64(10), pp.3395-3401.
 12. Barrett, C.F. and Parker, M.A., 2005. Prevalence of *Burkholderia* sp. nodule symbionts on four mimosoid legumes from Barro Colorado Island, Panama. *Systematic and Applied Microbiology*, 28(1), pp.57-65.
 13. Santini, A.C., Santos, H.R., Gross, E. and Corrêa, R.X., 2013. Genetic diversity of *Burkholderia* (Proteobacteria) species from the Caatinga and Atlantic rainforest biomes in Bahia, Brazil. *Genetics and Molecular Research*, 12(1), pp.655-664.
 14. Moreira, F.M.S., 2006. Nitrogen-fixing Leguminosae-nodulating Bacteria. *Soil biodiversity in Amazonian and other Brazilian ecosystems*, p.237.